Mathematical Modeling of the Calcium Binding Proteins Calmodulin and Troponin C

Undergraduate Research Thesis

Presented in partial fulfillment of the requirements for graduation *with honors research distinction* in the undergraduate colleges of The Ohio State University

by

Garrett Hauck

The Ohio State University

2021

Project Advisor: Professor Jonathan Davis, Department of Physiology and Cell Biology

Abstract

Cellular calcium either directly controls, or at a minimum modulates nearly, if not, all human functions. As such, the structural motif that forms the calcium binding pocket (EF-hand) is one of the most common protein elements found in Mammalia. The calcium binding properties of the EF-hands appear to be "tuned" to receive and respond to different calcium signals and can vary by over 6 orders of magnitude. Our laboratory is attempting to understand the rules that govern calcium binding to proteins, so that we might be able to engineer these proteins as gene therapies to treat diseases. Calmodulin is a small, switch-like, calcium binding protein that plays multiple roles in all human cells depending on what protein system it is attached. We are using rational design, super computers as well as evolutionary clues to help us engineer calmodulins with specific functionalities. Using a simple mathematical model to describe the calcium binding events of calmodulin and available steady-state and kinetic data, we determined the exact rate parameters that govern this interaction for wild type calmodulin and other variants (such as plant calmodulins, disease mutants, and engineered calmodulins). Troponin C is the "cousin" of calmodulin and is very important in the calcium regulation of muscle contraction. Using a previously created mathematical model for the activation of the cardiac troponin complex, we modified the rate parameters to effectively simulate the calcium binding data for the slow skeletal isoform and characterized the effects of a potential small molecule therapeutic. This information better helps us identify what we want to alter in order to achieve a particular outcome (in terms of engineering novel calmodulins) and may help explain the different responses between cardiac and slow skeletal troponin.

Table of Contents

Chapter 1

In nearly every branch of science, from the physical sciences to the social sciences, the concept of modeling is incredibly important. Whether the subject is molecular biology, quantum physics, abnormal psychology, or criminology, a model will always be of use to aid in the description of systems and their study. In addition to simplifying a system of interest, models can be used to perform experiments that may not be feasible for a number of reasons – too costly, too risky, or too difficult. Researchers or others may also opt to use a model system to anticipate the impacts of an action or experimental modification. Perhaps a demolition crew would like to model the impact of their explosives on a building in an empty field before tearing down a building in the heart of a city, or perhaps a biochemist would like to anticipate whether an experiment with costly reagents is worth pursuing before spending the money. In any sense, a model can be beneficial in practically any scenario, and it is clear that models for biochemical systems can aid in the understanding of their molecular interactions.

When it comes to modeling biochemical interactions, there are several approaches that can be taken. One of the most frequently used approaches is modeling a system with molecular structures. A molecular structure can be obtained in many different ways – there are the experimental approaches of X-ray crystallography, nuclear magnetic resonance, and cryogenic electron microscopy, and a structure can also be predicted using the plethora of homology modeling software that now exists. If the structure of a protein or nucleic acid is known or predicted, researchers can use the model of the macromolecule to make predictions about interactions, investigate the potential impacts of mutations, and visualize potential or real conformational changes. These structural models are very powerful, allowing the scientist to view an extraordinarily small system on a computer screen. In addition to structural models, models can

be made from a mathematical standpoint to simulate the kinetics of a system, which can be used to describe binding events in a different fashion than structural models. The advantages of this approach are plentiful and distinct from a molecular structure. Firstly, the system of interest can be expanded to include more than a structural model. A mathematical model can involve several macromolecules and/or several small molecules or ions and can simulate several different binding events. Second, a mathematical model can allow other questions to be asked or other aspects of a system to be looked at. A mathematical model uses microscopic properties of individual molecules to describe binding events that cannot be investigated through structure. The affinity for a target cannot be determined by looking at a structure (although it can be estimated or predicted relative to similar interactions), but it can be determined using the specifics of a mathematical model. Finally, mathematical models are more efficient than most experiments that involve structural models. The software required is cheaper, there is less demand in terms of computing power, and the experiments simply do not consumer as much time. With the same resources, more work can be accomplished in one day with a mathematical model than most experiments that a structural model is used for.

It is this concept of mathematical modeling that this project is centered around. In the context of the calcium binding protein calmodulin (CaM), the protein-ion binding events are useful to model to obtain a deeper understanding of the microscopic mechanisms that govern the properties of this interaction. Similarly, the calcium binding protein troponin C (TnC), part of the troponin complex (Tn), and its interaction with troponin I (TnI), can be effectively modeled to understand the specifics of the binding events. Both of these proteins are important for function and regulation of the contraction/relaxation cycle in muscle, and hence have serious implications

in several disease states, such as arrythmia and various cardiomyopathies. For these reasons, it is important to understand them with as much depth as possible.

Properties of Calmodulin

As a protein, CaM is arguably one of the most influential proteins to be characterized and researched. CaM is essential for mammalian life, so much so that there are multiple copies of the CaM gene in certain organisms to prevent lethal recombination events involving the CaM gene (Fischer et al., 1988). CaM is a calcium-binding protein that allows for signals communicated by changing calcium levels to be translated into specific cellular and biochemical events (Chin & Means, 2000). CaM has hundreds of downstream binding targets, which can include a plethora of different protein classes, examples including Myosin Light Chain Kinase (MLCK), Calcineurin (CaN), Nitric Oxide Synthase (NOS), and the Ryanodine Receptor (RyR) (Vogel, 1994).

CaM is a small protein of 148 amino acid residues, consisting of two domains which create a dumbbell shape. The two domains are connected via a flexible linker, allowing for a large amount of flexibility within the CaM structure, hence providing the opportunity for the vast array of downstream targets for CaM. Each domain contains two calcium binding sites, which are part of the EF-hand family, providing four binding motifs in one protein. The EF-hand structural motif consists of a helix-loop-helix pattern, where the loop contains amino acids designed to coordinate a calcium ion. In addition to binding calcium, CaM is known to competitively bind magnesium ions, competing with up to 1 mM free magnesium in the cell (Ohki et al., 1997; Tikunova et al., 2001; Waters, 2011). In general, the magnesium affinity of the C-terminal domain of CaM is physiologically irrelevant ($K_D > 10$ mM), but the affinity for magnesium of the N-terminal domain of CaM is physiologically relevant, approximately 1 mM (Walton et al., 2017).

When CaM binds calcium, it undergoes a conformational change. In the apo state, the surface of CaM is hydrophilic and hydrophobic residues are buried, but upon calcium binding, hydrophobic pockets are exposed within both the N-terminal and C-terminal domains, which will then partake in target binding (Chin & Means, 2000). Interestingly, the CaM-binding domains of the many CaM targets do not share sequence similarity. However, this can be a rationalization for why CaM can have so many protein targets. The general formula for a CaM-binding domain includes basic and hydrophobic residues arranged in an amphipathic alpha helix. This allows for CaM's methionine-rich interior to favorably interact with the targe enzyme and alter function within the cell (Vogel, 1994). Evidently, there are many aspects of CaM that can be and deserve to be studied, many of which can be aided by the utilization of models.

While the C-terminal and N-terminal domains of CaM share structural similarity, their calcium binding properties are different. In terms of both steady-state affinity and kinetics, the two domains will bind calcium differently, which can help determine a potential method of regulation through CaM acting as a calcium sensing switch. For steady-state affinity, both domains have calcium affinities in the micromolar range, but the C-terminal domain is more sensitive to calcium than the N-terminal domain by approximately 2-3-fold (Walton et al., 2017). Interestingly, the kinetics of calcium binding to these domains do not correlate to the affinity difference in all cases. The calcium association rate is faster for the N-terminal domain by 1-2 orders of magnitude, which would imply a higher calcium affinity for the N-terminal domain (Johnson et al., 1996). However, the higher affinity of the C-terminal domain comes from the difference in calcium dissociation rates. The calcium dissociation rate of the N-terminal domain is approximately 2 orders of magnitude faster than the calcium dissociation rate of the C-terminal domain, so much so that the overall affinity of the C-terminal domain is higher than that of the N-terminal domain (Johnson et

al., 1996). These data imply a potential method of how CaM can act as a calcium sensing switch and activate target enzymes based on calcium signals. The affinity of the C-terminal domain for calcium is low enough that a significant level of CaM C-terminal domains will be saturated at basal calcium levels, hence the C-terminal domain will be in an open conformation and able to bind target enzymes. On the other hand, the N-terminal domain's affinity is too low to hold on to calcium at a basal level, but its fast calcium association rate allows for rapid calcium binding followed by rapid target enzyme binding and subsequent activation. In essence, the C-terminal domain acts an anchor to the target enzyme, and the N-terminal domain acts as the regulator to activate the target enzyme when the signal is given (Davis et al., 2016).

Calmodulin Variants

Much of the current literature in the field focuses on CaM's relevance to human health and disease, but this protein is not limited to humans, or even animals. While it has been established that there is essentially no variance among CaM sequence and structure among vertebrates, plant species express many different variants of CaM with differences that can range from a single amino acid up to a quarter of the entire protein (McCormack and Braam, 2003; McCormack et al., 2005). In addition to the structural variation present between the CaM isoforms found in plants, there is also substantial variation in the expression patterns between certain isoforms (Al-Quraan, 2008; Park et al., 2009). When it comes to CaM, one of the better characterized plant organisms is the soybean. There are two main isoforms in the soybean: soybean CaM 1 (sCaM1) and soybean CaM 4 (sCaM4). sCaM1 is known as the "housekeeping" CaM, expressed during periods of low stress where the plant is functioning as desired. sCaM4 becomes expressed during periods of stress and exhibits differential effects from sCaM1. These two isoforms will bind nearly all the same target proteins, but sCaM4 inhibits a group of enzymes and ion channels that sCaM1 activates (Lee et

al., 1995, 2000; Cho et al., 1998; Kondo et al., 1999). Interestingly, this phenomenon of isoform-dependent activation and inhibition is not limited to plant targets. When the soybean CaMs bind to human CaN, sCaM1 will activate the phosphatase to a similar level as mammalian CaM, but sCaM4's activation of CaN is decreased compared to the mammalian isoform (Cho et al., 1998).

The soybean CaMs have more differences between them than their effects on target binding. Consistent with mammalian CaM, the C-terminal domain of both soybean isoforms bind calcium with approximately four-fold higher affinity than the N-terminal domain (Walton et al., 2017). In addition, each domain of sCaM4 has a higher affinity for calcium than the domains of sCaM1 (Walton et al., 2017). In terms of magnesium affinity, the soybean CaMs follow the same pattern as mammalian CaM, with a physiologically relevant N-terminal domain affinity and a physiologically irrelevant C-terminal domain affinity (Walton et al., 2017). Between the two isoforms there is a similar pattern to the calcium affinity, with sCaM4 having a higher affinity for magnesium than sCaM1, however both soybean isoforms have a higher ion affinity than mammalian CaM (Walton et al., 2017). In addition to the steady-state differences, there are also differences in the kinetics of these two CaMs. The calcium dissociation rates from each domain are slower for sCaM4, corresponding to the higher calcium affinity shown by sCaM4. Interestingly, the calcium association rates are faster for sCaM1 even though this isoform has a lower steady-state affinity (Walton et al., 2017). This interplay between steady-state affinity and kinetics is thought-provoking and perhaps demonstrates different methods of regulation by these CaMs or insights to the evolutionary driving forces behind their divergence. Comparing the C-terminal and N-terminal domains in the individual proteins, the pattern of their kinetics is similar to mammalian CaM. The N-terminal domain of each soybean CaM has a much higher calcium dissociation rate than the corresponding C-terminal domain (Walton et al., 2017). However, there

is no available data for the calcium association rates of the C-terminal domains of the soybean CaMs, so these rates can only be compared to the calcium association rates of the respective N-terminal domains if they are calculated based on the K_D and the calcium dissociation rate.

In addition to CaMs in other organisms, it is also interesting to look at artificial variants of CaM. One of the primary efforts of the lab is to engineer novel CaMs to bind both calcium and target proteins with higher affinity, with the end goal to use a successfully engineered protein as treatment for cardiovascular disease. Through *in silico* protein redesign and molecular dynamics simulations, mutations can be pinpointed in CaM that will enhance its efficacy within a certain system. Once purified, the engineered CaMs go through a series of biochemical and physiological experiments to test whether they act as predicted. The biochemical characterizations, such as determining the steady-state calcium affinity and obtaining the kinetics of the engineered CaM, are more fundamental and aid in future protein redesign. The physiological characterizations more directly establish whether the engineered protein could potentially be therapeutic. The lab has made progress on this front, using clues from previous engineered proteins as well as evolutionary hints from the soybean CaMs to engineer several CaMs with potential to improve the functioning of the biochemical systems that go awry in disease. The current focus is on engineering a CaM to bind RyR with high affinity (Bogdanov et al., 2020).

Previous Calmodulin Models

This project is not the first to approach CaM binding events from a mathematical modeling standpoint. Other researchers in the field have previously used mathematical models to simulate the effects of competition on the differential activation of different CaM downstream binding targets. The experiments performed with these models are useful and answer important questions. The models can be used to demonstrate some of the shortcomings of modeling CaM binding to

target proteins in isolation, specifically that competition for CaM as a limiting factor with many target proteins influences the activation levels of individual targets (Romano et al., 2017). A model can also demonstrate the impact of the frequency of calcium oscillations on differential activation of multiple CaM binding targets. Depending on the presence or absence of competitive target proteins, the activation of specific target proteins, such as CaN or NOS, will be maximal at differing frequencies of calcium transients, and the range of optimal frequencies will sharpen (Romano et al., 2017; Slavov, Carey, & Linse, 2013). These models also show how the optimal oscillation frequency can change given the number of competitors (Slavov, Carey, & Linse, 2013).

In addition to demonstrating frequency dependence of differential activation of CaM target proteins, models have been used to show how biochemical systems maintain homeostasis in times of stress and perturbation. In addition to changes in the activation of different CaM target proteins based on the frequency of calcium oscillations, the same effect will be seen if other parameters of the system are modified (Romano et al., 2017). Given the observations from these experiments on "competitive tuning," there are a multitude of potential consequences, including some that can be particularly relevant to physiology. In the context of a neuron, the most relevant downstream CaM binding proteins are involved in neuronal plasticity, and any slight perturbation in one of these systems has the potential to have dramatic impacts on neurophysiology. Using a mathematical model to describe the calcium binding events of CaM as well as the target protein binding events, it has been demonstrated that the competitive tuning that is seen between the target proteins is used to provide a compensatory mechanism for perturbations in the system (Pharris, Patel, & Kinzer-Ursem, 2018). If a CaM target protein is downregulated or knocked out, the system can shift in such a way that, through competitive tuning, a separate CaM target protein is upregulated that has

a similar function that leads to the same final outcome as before (Pharris, Patel, & Kinzer-Ursem, 2018).

Evidently, the CaM models in the literature are complex. The incorporation of just one target protein into a model can dramatically increase the number of parameters necessary for the simulation, yet most of the previous models incorporate multiple. However, with increasing complexity also comes simplicity. When the parameter field becomes large with many reactions taking place in the simulation, researchers will look for ways to simplify the model in any way that they can, and this usually is done with regard to the calcium binding events of CaM. As previously stated, each molecule of CaM has four calcium binding sites, which allows for 16 potential configurations of the individual sites being bound or unbound, ranging from all unbound to all bound. But there are some ways of cutting down on this breadth of possibilities. Firstly, the assumption can be made that there is no significant difference between the individual sites within each domain. Without the need to distinguish between site 1 and site 2 within a domain, the model for calcium binding can be reduced to nine potential configurations (Pepke et al., 2010). This model can be further simplified to create a four-state model by ignoring the configurations in which only one calcium ion is bound to a domain, only considering the apo state, the calcium saturated state, and the states in which one domain is saturated and the other is apo. The rationalization behind condensing the individual calcium binding events into a single event comes from the highly cooperative nature of the calcium binding (Pepke et al., 2010). To simplify the model to three states, a model can be operated under the assumption that the C-terminal domain will always be saturated before the N-terminal domain, so the configuration in which the N-terminal domain is saturated, and the C-terminal domain is unsaturated is neglected (Slavov, Carey, & Linse, 2013). However, this assumption is based solely on the steady-state data which indicates that the C-

terminal domain has a higher calcium affinity than the N-terminal domain (Walton et al., 2017) but neglects to consider that the calcium association rate is much faster for the N-terminal domain than the C-terminal domain (Johnson et al., 1996). There is also the possibility of condensing the model even further to only two states, an all or nothing model in which CaM is either apo or fully saturated with calcium (Pharris, Patel, & Kinzer-Ursem, 2018).

Of course, this simplification does not occur without consequence. Clearly, a model with more possible states is more biophysically accurate. Comparing the two simplest models, a four-state model is more sensitive to calcium than a two-state model is (Pharris, Patel, & Kinzer-Ursem, 2018), indicating that a significant level of biophysical accuracy is lost when the model is oversimplified to this degree. Nevertheless, it has been argued that the four-state model has negligible differences between the nine-state model (Romano et al., 2017), hence a three-state or four-state model has generally been considered to be adequate to simulate the calcium binding properties of CaM within the larger-scale model that includes multiple target proteins (Romano et al., 2017; Pharris, Patel, & Kinzer-Ursem, 2018; Slavov, Carey, & Linse, 2013).

However, this simple model may not be accurate enough to effectively model the calcium binding in all scenarios. Returning to the idea of cooperativity that was used to simplify the model from nine to four states, the cooperativeness of the calcium binding could end up being an issue in certain modeling experiments. The nine-state model will take into account the cooperativity of the calcium ions, but a four-state model will not. In the experiments previously described in which the calcium concentration oscillates between saturating and subsaturating, the cooperativity between calcium ions will not have a dramatic effect, as the free calcium concentration is never such that CaM will be partially saturated, so the four-state model will be sufficient. But if simulations were to be run in which the free calcium concentration would linger at a value in which CaM would be

partially saturated, it is important that cooperativity is considered, because a cooperative model will give a level of calcium saturation that can be significantly different than an uncooperative model.

Due to this situation in which cooperativity can and will have an impact on the level of calcium saturation, it is important to have a model that both exhibits the cooperativity required and accurately describes the calcium binding events for CaM. This is what the current project aims to accomplish. The creation of an accurate and descriptive model can be extraordinarily beneficial to the field. It can be incorporated into more complex models that describe CaM's interactions with target proteins, specifically creating a model that can be accurate during subsaturating calcium concentrations. It can also be used in the quest to engineer novel CaMs as potential therapeutics, potentially allowing for a deeper understanding of the effects certain mutations have on the kinetics of calcium binding. With a model that contains a sufficient number of configurations and that is accurate, many doors are opened to new experiments, new models, and new information.

Troponin C and the Troponin Complex

TnC is similar to CaM in both structure and in function. Like CaM, TnC belongs to the EF-hand superfamily of proteins and consists of two domains, N-terminal and C-terminal, connected by a flexible linker. Each of the two domains contains two EF-hand motifs that can coordinate and bind calcium ions (and magnesium ions). The C-terminal domain of TnC has a high affinity for both calcium and magnesium and is most often implicated in complexing TnC to the other troponin subunits – TnI and troponin T (TnT) (Davis & Tikunova, 2008; Jin, 2014). In the isoforms that are discussed in this thesis, the N-terminal domain of TnC the first EF-hand motif is unable to bind calcium, but the second motif can and will bind calcium and magnesium, but with a lower affinity

than the C-terminal domain (Tikunova & Davis, 2004). Similarly to CaM, this allows for the C-terminal domain to act as an anchor and for the N-terminal domain to play the regulatory role.

This regulatory role allows for TnC to be a calcium-sensing switch to regulate contraction and relaxation in both cardiac and skeletal muscle. Once calcium binds the N-terminal domain of TnC, contraction is initiated. Through a series of conformational changes, the switch peptide region of TnI binds TnC, the inhibitory peptide region of TnI releases from actin and leads to repositioning of tropomyosin, leaving the myosin binding sites on actin exposed. The resulting interactions between actin and myosin produce the force that is seen during muscle contraction. After the stimulus ceases, the dissociation of calcium from TnC initiates the relaxation process. The inhibitory region of TnI once again binds actin, tropomyosin is again repositioned back to its inhibitory spot, and the actin-myosin interactions are blocked (Layland et al., 2005; Davis & Tikunova, 2008). Due to its role in the critical physiological process of muscle contraction and relaxation, the significance of TnC cannot be understated. Mutations in TnC, as well as the other proteins in the troponin complex, have been linked to cardiomyopathy (Liu et al., 2012a; Liu et al., 2012b; Li et al., 2013). Specifically, mutations that enhance apparent calcium affinity of troponin are linked to hypertrophic and restrictive cardiomyopathy (Davis et al., 2007a; Liu et al., 2012b; Li et al., 2013; Alves et al., 2014), and mutations that decrease the apparent calcium sensitivity are implicated in dilated cardiomyopathy (Liu et al., 2012b). However, there are several different ways that apparent calcium sensitivity can be modified. As such, it is unlikely that each individual mechanism would lead to a disease state. For example, a calcium sensitized TnC mutant has been shown to improve contraction and cardiac function in mouse models of myocardial infarction and does not lead to disease in healthy mice (Shettigar et al., 2016). Evidently, TnC has

an important role in muscle physiology, and it is important to understand the mechanisms by which its calcium binding and structural dynamics within the troponin complex are regulated.

Interestingly, the calcium binding properties of TnC depend on its interactions with other troponin proteins as well as thin and thick filament proteins (Davis et al., 2007b), similar to CaM and its target proteins. The N-terminal domain of the isolated TnC protein has a relatively low calcium binding affinity and a fast rate of calcium dissociation (Tikunova & Davis, 2004). Complexing TnC with TnI and TnT increases the affinity for calcium and decreases the calcium dissociation rate (Davis et al., 2007b; Liu et al., 2012b). In reconstituted thin filaments containing S1 myosin, the calcium affinity is highest, and the calcium dissociation rate is slowest (Tikunova & Davis, 2004; Little, 2012). However, in reconstituted thin filaments alone, the calcium affinity is lower than the troponin complex and thin filaments with S1 myosin (although not as low as the isolated protein), and the calcium dissociation rate shows the same pattern. Given that TnI binding to TnC has a large effect on stabilization, it makes sense that the troponin complex and the thin filaments have higher calcium affinities than the isolated TnC protein. But the lower calcium affinity and higher calcium dissociation rate in the thin filament compared the troponin complex and the thin filament with S1 myosin needs further explanation. To do this, the idea of effective concentration can be introduced. Within the troponin complex, TnC "sees" TnI at a high frequency because the proteins do not dissociate and operate in a limited volume, in essence increasing their effective concentration so they are free to interact, increasing the stability of TnC and hence increasing the calcium affinity (Siddiqui et al., 2016). In the thin filament, there is competition for TnI from actin and tropomyosin, so the effective concentration of TnI is lower and TnC's calcium affinity suffers as a result (Siddiqui et al., 2016). The addition of S1 myosin undoes much of the actin and tropomyosin competition, restoring the TnI effective concentration and hence the

calcium affinity and calcium dissociation rates of TnC to levels similar to the troponin complex (Siddiqui et al., 2016; Davis et al., 2007b).

In addition to mutations in TnC, mutations and post-translational modifications of the other troponin proteins can modulate the behavior of muscle and/or cause it to enter a disease state. Mutations in TnI and TnT have been implicated in forms of dilated and restrictive/hypertrophic cardiomyopathy (Du et al., 2007; Liu et al., 2012a; Liu et al., 2012b; Li et al., 2013). Mutations in TnI and TnT as well as post-translational modifications such as phosphorylation have previously been characterized in terms of their effects on the calcium binding properties of TnC, both as part of the troponin complex and as part of the thin filament (Liu et al., 2012b; Liu et al., 2014). The general pattern seen is that mutations associated with dilated cardiomyopathy decrease the calcium affinity of TnC in thin filaments, and mutations associated with hypertrophic and restrictive cardiomyopathy increase the calcium affinity (Liu et al., 2012b). Phosphorylations can also influence the calcium affinity of TnC in thin filaments, either increasing or decreasing the affinity based on the location of the modification (Liu et al., 2014). Interestingly, while these mutations and post-translational modifications have effects on the thin filament calcium sensitivity, the majority of them do not change the calcium binding properties of the isolated troponin complex, potentially indicating that the mutations and modifications affect how the troponin complex interacts with actin and tropomyosin as opposed to affecting the troponin complex itself (Liu et al., 2012b; Liu et al., 2014).

It is important to note that the troponin complex, and hence the contractile apparatus as a whole, is not completely the same between different muscle types. Within the three muscle types that contain troponin – cardiac, slow skeletal, and fast skeletal (there is no troponin present in smooth muscle), each have their own isoform of each troponin protein. The only exception is with

TnC, where cardiac and slow skeletal muscle share the same TnC isoform. Evidently, the calcium binding properties of the troponin complex and of the thin filament differ between muscle types. For example, the slow skeletal troponin complex is more sensitive to calcium than the cardiac troponin complex by ~3-fold, and the calcium dissociation rate for the slow skeletal troponin complex is slower by a similar magnitude (Davis et al., 2007b; Gomes et al., 2004). These differences are less dramatic when considering the reconstituted thin filament, but are still seen nonetheless (Davis et al., 2007b; Gomes et al., 2004). These comparisons are intriguing because the TnC isoform is unchanged between cardiac and slow skeletal muscle, so any differences in calcium binding properties must come from differences in the non-calcium binding proteins present in the troponin complex (TnI and/or TnT). It is likely that TnI differences exert stronger influences, given that TnI and TnC interact as part of the contractile process.

Previous Troponin Models

In addition to mathematical models of CaM, this project is not the first to describe the calcium binding properties of the troponin complex using a mathematical model. The Davis lab has previously utilized this approach to describe the calcium binding properties of the troponin complex and the thin filament with the cardiac isoforms (Siddiqui et al., 2016). The model has four states: (1) TnC, (2) TnC-Ca^{2+}, (3) TnC-Ca^{2+}-TnI, and (4) TnC-TnI (Figure 1). The requirement for four states can be explained by building up from fewer states to more states. A two-state model with only TnC and TnC-Ca^{2+} can simulate the steady-state and kinetic properties of the troponin complex and the thin filament, but this model then must assume that the intrinsic calcium association and dissociation rates from TnC change in order to explain the observed differences between systems. This would implicate that within different systems the EF-hand in TnC changes structure, the calcium coordination changes, or the way TnC interacts with TnI and/or

TnT changes, which is not as logical as changes in calcium binding properties being due to particular perturbations that do not include structural changes in calcium-bound TnC, such as the binding of TnI (Siddiqui et al., 2016). Thus, it is logical that a third state should be added in which TnC is bound to both calcium and TnI. With the addition of TnI to the model, the differences in calcium affinity and dissociation rate can be simulated with the same intrinsic rate parameters solely by altering the effective concentration of TnI (Siddiqui et al., 2016). However, the three-state model predicts that as TnI effective concentration is increased, the affinity of the system for calcium approaches infinity (K_D approaches 0) and the calcium dissociation rate asymptotes at 0, which is not the case based on biochemical studies (Siddiqui et al., 2016). Therefore, the fourth state is required, created when the TnC-Ca^{2+}-TnI state dissociates by releasing calcium as opposed to TnI. With the four-state model, an infinite increase in TnI effective concentration leads to the calcium affinity and dissociation rate approaching the values that are measured for the troponin complex, in which a saturating effective concentration of TnI can be assumed (Siddiqui et al., 2016). Thus, a four-state model is deemed to be accurate for describing the calcium binding properties of TnC within various systems.

However, since there are different isoforms of the troponin proteins in the different muscle types, this previous model is only valid for cardiac muscle since the experimental verification used was data from the cardiac isoforms. Of course, there is nothing wrong with centering attention at cardiac muscle – any defects in cardiac contractility will have severe effects on an individual. But there are also merits to studying other muscle types, both to aid in understanding disease states in noncardiac muscle and potentially using insights from noncardiac muscle to aid in the effort to therapeutically modify cardiac muscle. For these reasons, a mathematical model which describes

the calcium binding properties of noncardiac muscle types would be useful. This is another goal of the current project.

In addition, a mathematical model for troponin can be useful to learn more about the effect of potential small molecule drugs on the system. As described earlier, changing the properties of troponin through mutation or post-translational modification can have a dramatic effect on the calcium binding properties of the system and hence the contractile properties, so it is logical to presume that a small molecule that can change the calcium binding properties of TnC has potential as a therapeutic. Using experimental data to obtain the affinity and kinetics of small molecule binding as well as which specific state binding occurs in, the small molecule can be incorporated into the mathematical model and the details of its binding and/or potential implications of further tweaking of the system can be investigated. In essence, a previously built troponin model is effective at simulating the calcium binding properties of TnC in different systems, but this project aims to expand that model to different troponin isoforms and to incorporate a potential small molecule therapeutic into the model.

Chapter 2 - Methods

Although this thesis will discuss experiments performed with multiple CaM variants as well as various troponin protein isoforms, a significant portion of the experimental (wet lab) data collected with these proteins comes from work performed before the onset of this thesis and by researchers other than the author, so for the sake of brevity only the most relevant wet lab experiments and the proteins used in these experiments will be elaborated upon.

Protein Expression & Purification

Mammalian wild type CaM, sCaM1, sCaM4, and TnC DNA constructs in the pET17b vector were expressed in BL21 DE3 *Escherichia coli* in LB medium, with DNA expression aided by the presence of IPTG to activate the promoter. The DNA construct for the engineered CaM as well as the constructs for CaM^{F19W}, $sCaM1^{F19W}$, $sCaM4^{F19W}$, TnC^{F27W}, and TnC^{T53C} were generated from their respective wild type plasmids. Primer-selected site-directed mutagenesis was performed using a QuikChange Lightning Multi-site kit from Agilent following the manufacturer's PCR protocol. Following PCR, the identities of the engineered CaM, the CaM mutants, and the TnC mutants were confirmed by DNA sequence analysis (GeneWiz), and subsequently expressed in the same manner as the other proteins. Following DNA expression in the bacteria, the various proteins were purified using standard laboratory procedures including sonication, centrifugation, precipitation, and column chromatography.

Steady-State Calcium Titrations

To measure the steady-state calcium affinity of the different CaM variants, steady-state calcium titrations were performed. All steady-state fluorescence measurements were performed using a Perkin-Elmer LS 55 spectrofluorometer at 20 ℃, with the temperature maintained by a circulating water bath. The titration buffer consisted of 200 mM 3-(N-morpholino)propanesulfonic acid (MOPS), 150 mM KCl, and 2 mM ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid (EGTA, a calcium chelator) at pH 7.0. To follow the calcium dependent change in the C-terminal domain of CaM, Tyr fluorescence was excited at 275nm and monitored at 305nm as microliter amounts of $CaCl_2$ were added to 2 ml of each CaM (1 μM) in titration buffer with constant stirring. The free calcium concentration was calculated using the EGCA02 program

(Robertson & Potter, 1984). To follow the calcium dependent change in the N-terminal domain of CaM, there are two methods, both of which were used. The first option is to mutate the very faintly fluorescent Phe19 residue in CaM (and both soybean CaMs) to a more prominently fluorescent Trp residue. In this case, Trp fluorescence was excited at 295nm and monitored at 350nm as microliter amounts of $CaCl_2$ were added to the reaction mixture as previously described. The second option is to use an extrinsic dye such as the fluorescent hydrophobic dye 4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid (bis-ANS) to monitor the change, as the dye fluoresces in a hydrophobic environment, such as when the N-terminal domain binds calcium to expose its hydrophobic core. In this case, bis-ANS fluorescence was excited at 390nm and monitored at 500nm as microliter amounts of $CaCl_2$ were added to the reaction mixture as previously described.

The steady-state calcium affinity of TnC within different systems was measured with a similar procedure (Siddiqui et al., 2016). When monitoring calcium dependent conformational changes in TnC, only the N-terminal domain must be considered. Similar to the CaMs, this domain has no intrinsically fluorescent amino acids other than a faintly fluorescent Phe residue at position 27. Thus, to follow the calcium dependent change in TnC (in any system), Phe27 was mutated to a more prominently fluorescent Trp residue, and the steady-state calcium titrations were performed as described above, except at 15 °C. Each reported experimentally determined K_D represents an average of at least three successive titrations ± standard error and fit with the Hill equation.

Kinetic Experiments

All kinetic data for CaM were collected using an Applied Photophysics Ltd. Model SX.20 MV stopped-flow instrument with a dead time of 1.4 ms at 20 °C. Kinetic data for troponin systems were collected using an Applied Photophysics Ltd. Model SX.20 stopped-flow instrument with a dead time of 1.4 ms at 15 °C. The buffer for all stopped-flow experiments consisted of 10 mM

MOPS and 150 mM KCl at pH 7.0. When necessary, Tyr fluorescence (used to follow calcium dependent changes in the C-terminal domain of CaM) was excited at 295nm and monitored using a UV-transmitting black glass (UG1) filter from Oriel, Trp fluorescence (used to follow calcium dependent changes in the N-terminal domain of CaM) was excited at 295nm and monitored using the same UG1 filter, bis-ANS fluorescence (used to follow the calcium dependent changes in the N-terminal domain of CaM as well) was excited at 390nm and monitored using a 500nm long pass interference filter from Newport (Walton et al., 2017). To follow calcium dependent changes in TnC in various systems, instead of using the TnC^{F27W} protein, a different approach was taken. A T53C mutation was introduced, and the new Cys residue was subsequently labelled with 2-(4′-(iodoacetamido)anilino)naphthalene-6-sulfonic acid (IAANS), an environmentally sensitive fluorophore with a good signal-to-noise ratio. IAANS fluorescence was excited at 330nm and monitored using a 420-470nm band pass interference filter from Oriel (Siddiqui et al., 2016).

The first set of kinetic experiments are calcium dissociation experiments. To measure the calcium dissociation rate from the C-terminal domain of CaM, stopped-flow buffer with 200 μM calcium and 3 μM CaM was rapidly mixed with 10 mM EGTA in stopped-flow buffer (Walton et al., 2017). To measure the calcium dissociation rate from the N-terminal domain of CaM, stopped-flow buffer with 200 μM calcium and 1 μM CaM^{F19W} (or 1 μM CaM and 0.5 μM bis-ANS) was rapidly mixed with 10 mM EGTA in stopped-flow buffer (Walton et al., 2017). These experiments were performed with all CaM variants. Calcium dissociation rates from TnC were measured in a similar fashion (Siddiqui et al., 2016). As an example, for isolated TnC, stopped-flow buffer with saturating calcium and 1 μM T53C-IAANS TnC was rapidly mixed with 10 mM EGTA in stopped-flow buffer (Siddiqui et al., 2016).

Next, there are calcium association experiments. While these do not measure intrinsic calcium association rates (which are usually in the range of $10^6 – 10^9$ $M^{-1}s^{-1}$), they do give apparent rates (usually ~300/s or less) that change with varying initial calcium concentrations. To measure the calcium association rate to the C-terminal domain of CaM, stopped-flow buffer with 3 μM CaM and 15 μM EGTA was rapidly mixed with increasing concentrations of calcium (2.5, 5, 10, 20, and 40 μM) in stopped-flow buffer. There were no experiments performed to measure the calcium association rate to the N-terminal domain of CaM, and the experiments described with the C-terminal domain were only performed with mammalian CaM and the engineered CaM. There were no experiments performed to measure the calcium association rate to TnC.

Finally, there are transient occupancy (TO) experiments. These indirectly measure calcium association rates by creating artificial calcium transients and monitoring the level of calcium saturation of CaM and TnC over time. To measure the TO of the N-terminal domain of CaM, stopped-flow buffer with 2 μM CaM^{F19W} and 2 mM EGTA was rapidly mixed with either 100 or 250 μM calcium in stopped-flow buffer (Walton et al., 2017). TOs were only performed with the two soybean CaMs, and there were no TO experiments performed following the C-terminal domain. There were no TO experiments performed to following TnC in any system that ended up being relevant to this project.

Model Creation & Simulations

All mathematical models were built and all simulations performed using the Copasi® software (Version 4.24). This software program is a standard biochemical modeling program which is user-friendly and very powerful. The program performs simulations by solving differential equations that are based on the inputted reactions that can take place, the inputted

kinetic parameters, and the concentration of each species at a given time point. For CaM models, the differential equations necessary are:

$$\frac{d[Ca]}{dt} = k_{off,1}[Ca_1CaM] + k_{off,2}[Ca_2CaM] + k_{off,EGTA}[Ca \cdot EGTA] - k_{on,1}[Ca][CaM] - k_{on,2}[Ca][Ca_1CaM] - k_{on,EGTA}[Ca][EGTA]$$

$$\frac{d[CaM]}{dt} = k_{off,1}[Ca_1CaM] - k_{on,1}[Ca][CaM]$$

$$\frac{d[Ca_1CaM]}{dt} = k_{on,1}[Ca][CaM] + k_{off,2}[Ca_2CaM] - k_{off,1}[Ca_1CaM] - k_{on,2}[Ca][Ca_1CaM]$$

$$\frac{d[Ca_2CaM]}{dt} = k_{on,2}[Ca][Ca_1CaM] - k_{off,2}[Ca_2CaM]$$

$$\frac{d[EGTA]}{dt} = k_{off,EGTA}[Ca \cdot EGTA] - k_{on,EGTA}[Ca][EGTA]$$

$$\frac{d[Ca \cdot EGTA]}{dt} = k_{on,EGTA}[Ca][EGTA] - k_{off,EGTA}[Ca \cdot EGTA]$$

where *CaM* is an apo domain of CaM (either N-terminal or C-terminal), Ca_1CaM is a CaM domain with one calcium ion bound, Ca_2CaM is a CaM domain with two calcium ions bound, and $Ca \cdot EGTA$ is calcium chelated by EGTA. Rate parameters with a subscript "1" describe the kinetic parameters of the binding of the first calcium ion to an apo CaM domain, rate parameters with a subscript "2" describe the kinetic parameters of the binding of the second calcium ion to Ca_1CaM, and rate parameters with a subscript "EGTA" describe the kinetic parameters of the chelating of calcium by EGTA. For troponin models, the differential equations necessary are:

$$\frac{d[Ca]}{dt} = k_{off,1}[TnC \cdot Ca] + k_{off,3}[TnC \cdot Ca \cdot TnI] + k_{off,EGTA}[Ca \cdot EGTA] - k_{on,1}[Ca][TnC] - k_{on,3}[Ca][TnC \cdot TnI] - k_{on,EGTA}[Ca][EGTA]$$

$$\frac{d[TnC]}{dt} = k_{off,1}[TnC \cdot Ca] + k_{off,4}[TnC \cdot TnI] - k_{on,1}[Ca][TnC]$$

$$\frac{d[TnC \cdot Ca]}{dt} = k_{on,1}[Ca][TnC] + k_{off,2}[TnC \cdot Ca \cdot TnI] - k_{off,1}[TnC \cdot Ca] - k_{on,2}[TnC \cdot Ca][TnI]$$

$$\frac{d[TnC \cdot Ca \cdot TnI]}{dt} = k_{on,2}[TnC \cdot Ca][TnI] + k_{on,3}[Ca][TnC \cdot TnI] - k_{off,2}[TnC \cdot Ca \cdot TnI] - k_{off,3}[TnC \cdot Ca \cdot TnI]$$

$$\frac{d[TnC \cdot TnI]}{dt} = k_{off,3}[TnC \cdot Ca \cdot TnI] - k_{on,3}[Ca][TnC \cdot TnI] - k_{off,4}[TnC \cdot TnI]$$

$$\frac{d[EGTA]}{dt} = k_{off,EGTA}[Ca \cdot EGTA] - k_{on,EGTA}[Ca][EGTA]$$

$$\frac{d[Ca \cdot EGTA]}{dt} = k_{on,EGTA}[Ca][EGTA] - k_{off,EGTA}[Ca \cdot EGTA]$$

where *TnC* is apo TnC, *TnC·Ca* is TnC bound to calcium, *TnC·Ca·TnI* is TnC bound to both calcium and TnI, *TnC·TnI* is TnC bound to TnI, and *Ca·EGTA* is calcium chelated by EGTA. Rate parameters with a subscript "1" describe the kinetic parameters of the binding of calcium to TnC, rate parameters with a subscript "2" describe the kinetic parameters of the binding of TnI to TnC-Ca^{2+}, rate parameters with a subscript "3" describe the kinetic parameters of the binding of calcium to TnC-TnI, rate parameters with a subscript "4" describe the kinetic parameters of the dissociation of TnC-TnI into its constitutive proteins, and rate parameters with a subscript "EGTA" describe the kinetic parameters of the chelating of calcium by EGTA.

Copasi® has multiple different functions that can be used to accomplish multiple different tasks. The "steady-state" task is used to calculate what the equilibrium concentrations of all species

will be, given starting concentrations inputted by the user. The "time course" task is used to calculate the transient concentrations of all species will be over a given period of time, given starting concentrations inputted by the user. Within this task, both the length of the simulation and the number of intervals can be specified by the user. Importantly, the user can opt to have the concentration of a species remain fixed throughout a steady-state or time course simulation as opposed to letting the species change in concentration based on the reaction mixture. This option allows for useful flexibility within simulations. Finally, the "optimization" task is used to calculate an ideal (or optimal) starting concentration for a species or rate parameter to match a given output or simulation result desired by the user. The program runs a large number of simulations (specified by the user, usually anywhere ranging from 100,000 to 1,000,000) and varies the specified parameter between given bounds to find the optimal value of the parameter.

To simulate steady-state calcium binding curves with CaM, the "steady-state" task was used. The inputted starting concentrations were 0.5 μM CaM and the calcium concentration was fixed at different values ranging from 1 pM to 316 μM (pCa 9–3.5), and for each calcium concentration the equilibrium concentration of Ca_2CaM was plotted. For troponin models, no steady-state binding curves were simulated, only verification of K_D was required. To do this, the "steady-state" task was used, with inputted started concentrations of 1 μM TnC and either 1.22 mM or 12.3 μM TnI (depending on whether the system was the troponin complex or the thin filament, respectively), and the calcium concentration was fixed at the K_D value. To verify the K_D, the "steady-state" task was run and the equilibrium concentration of TnC-Ca^{2+}-TnI was verified to be ~ 0.5 μM.

To simulate calcium dissociation curves with CaM, the "time course" task was used. The inputted starting concentrations were 1 μM Ca_2CaM and the calcium concentration was fixed at 0

μM. The time course simulation was run for 0.4 s, with an interval length of 1 μs, and a plot was created showing the concentration of Ca_2CaM over time. To calculate the calcium dissociation rate, the following equation was used, since the calcium dissociation curve is a single exponential:

$$off\ rate = \frac{\ln(2)}{t_{1/2}}$$

where $t_{1/2}$ is the time point at which the Ca_2CaM concentration is half of its original value (in this case, 0.5 μM). The same "time course" task was used to simulate calcium dissociation curves from troponin. The inputted starting concentrations were 1 μM TnC-Ca^{2+}-TnI and either 1.22 mM or 12.3 μM TnI (depending on whether the system was the troponin complex or the thin filament, respectively), and the calcium concentration was fixed at 0 μM. The time course simulation was run for 0.2 s, with an interval length of 1 μs. To calculate the calcium dissociation rate, the same equation was used, with the exception that $t_{1/2}$ in this case is the time point at which the TnC-Ca^{2+}-TnI concentration is half its original value (in this case, 0.5 μM).

To simulate calcium association curves with CaM, the "time course" task was also used. The inputted starting concentrations were 3 μM CaM, 6 μM EGTA (but with kinetics that match the other domain that was not being simulated), and either 2.5, 5, 10, 20, or 40 μM calcium. The time course simulation was run for 0.4 s, with an interval length of 1 μs, and a plot was created showing the concentration of Ca_2CaM over time. To calculate the calcium association rate, the following equation was used:

$$on\ rate = \frac{\ln(2)}{t_{1/2}}$$

where $t_{1/2}$ is the time point at which the Ca_2CaM concentration is half of its maximum value (which in this case is also its equilibrium value and varies based on initial calcium concentration).

To simulate transient occupancy curves with CaM, the "time course" task was used. The inputted starting concentrations were 1 μM CaM, 1 mM EGTA (with regular EGTA kinetics), and either 50 or 125 μM calcium. The time course simulation was run for 0.1 s, with an interval length of 1 μs, and a plot was created showing the concentration of Ca_2CaM over time. The transient occupancy was calculated by dividing the concentration of Ca_2CaM at 1.5 ms (close to the peak) by the total concentration of CaM (1 μM). The same "time course" task was used to simulate transient occupancy curves with troponin. The inputted starting concentrations were 1 μM TnC, either 1.22 mM or 12.3 μM TnI (depending on whether the system was the troponin complex or the thin filament, respectively), 500 μM EGTA (with regular EGTA kinetics), and 50 μM calcium. The time course simulation was run for 0.1 s, with an interval length of 1 μs, and a plot was created showing the concentration of TnC-Ca^{2+}-TnI over time. The transient occupancy was calculated by dividing the concentration of TnC-Ca^{2+}-TnI at 3.5 ms (close to the peak) by the total concentration of TnC (1 μM).

Chapter 3 – Results

The first step to creating a model for calcium binding to CaM is to determine the individual steps that should be included. There are essentially two options to choose from when considering only one domain of CaM (so, two calcium ions): the first is what can be called a series model, in which the first step is calcium binding to apo CaM and the second step is calcium binding to Ca_1CaM, and the second is a parallel model in which the first calcium ion that binds has the option to bind to site 1 or site 2 on CaM, creating two branches and two potential pathways for complete CaM activation (Figure 2). For the sake of simplicity, the series model was chosen to simulate the calcium binding properties of wild type and other CaMs, although attempts to model the system with a parallel model will be discussed as well.

Wild Type Calmodulin – C Domain

The first model to be built, and the one on which the most time was spent, was the mathematical model for calcium binding to the C-terminal domain of mammalian CaM (also referred to as wild-type or wt CaM). Given the experimental steady-state and kinetic data, the rate parameters for each step of the reaction were determined that accurately simulated both a steady-state calcium titration and various kinetic curves (see Figure 3). The parameters have several important features and relationships that are necessary to point out. Primarily, the cooperativity that is seen within the binding of the two calcium ions is reflected in the model. Before plugging in parameters, the structure of the model itself introduces cooperativity simply with the fact that each calcium ion has its own step – if each step had identical rate parameters, a simulate titration curve would still have a steepness to it indicative of cooperativity. In order to add to and effectively simulate the cooperativity, $k_{on,2}$ is faster than $k_{on,1}$, and $k_{off,1}$ is faster than $k_{off,2}$. Second, the experimentally observed calcium dissociation rate is the value of $k_{off,2}$, indicating that $k_{off,1}$ has no effect on the calcium dissociation rate and will only influence the overall K_D of the system. Similarly, the parameter which has the greatest effect on the calcium association rate of the system is $k_{on,1}$, hence $k_{on,2}$ only has a significant influence on the overall K_D of the system.

Other Calmodulin Variants – C Domain

As discussed previously, different variants of CaM have different calcium binding properties. Therefore, a different model must be created for each CaM variant. However, a new model can be built that accurately describes the calcium binding properties of other CaMs without changing the structure of the model, and by only changing the kinetic parameters. As such, a model with rate parameters that accurately simulate the calcium affinity and kinetic properties of the C-terminal domain of one of the lab's engineered CaMs was created (Figure 4). Comparing this

model to the wild type model allows for several interesting insights. Firstly, since the engineered CaM has a calcium dissociation rate ~4-fold slower than wild type CaM, the rate parameter $k_{off,2}$ is also ~4-fold lower than the $k_{off,2}$ of the wild type CaM. Second, it is worth noting that although this decrease in $k_{off,2}$ would increase the calcium affinity of the engineered CaM, it would not do so to the same magnitude as is seen in the experimental data (almost an order of magnitude). Hence, $k_{on,1}$ must increase and $k_{off,1}$ must decrease, compared to wild type CaM, to allow for the engineered CaM model to accurately simulate the experimental data. Notably, the change in $k_{off,1}$ is more dramatic than the change in $k_{on,1}$. $k_{on,1}$ has a less dramatic change because, as mentioned previously, it is the major parameter that influences calcium association rates, for which there is not a large difference between wild type CaM and the engineered CaM.

In addition to the engineered CaM, models can be built for the C-terminal domain of both soybean CaMs, again using the same basic structure of the model and only changing the rate parameters. A model with rate parameters that accurately simulate the calcium affinity and the calcium dissociation rate of sCaM1 was created (Figure 5). sCaM1 has a calcium affinity slightly higher than that of wild type CaM, but a calcium dissociation rate that is ~4-fold faster. Thus, the decrease in affinity caused by an increased in $k_{off,2}$ had to be counteracted by decreasing $k_{off,1}$ and increasing both $k_{on,1}$ and $k_{on,2}$ to match the experimental K_D. A model was also created with rate parameters that accurately simulate the calcium affinity and the calcium dissociation rate of sCaM4 (Figure 6). sCaM4 has a calcium affinity ~3-fold higher than wild type CaM, and a similar calcium dissociation rate. Thus, the main changes in the rate parameters were a decrease in $k_{off,1}$ and increases in both $k_{on,1}$ and $k_{on,2}$, similar to sCaM1. It is important to note that for both of the soybean CaM C-terminal domain models, there was no available experimental data for calcium association rates, so the rate parameters presented here are only one of a large set of possible solutions to these

models (given only calcium affinity and dissociation rate to match), and it is possible that these rate parameters would not accurately simulate calcium associate rate or transient occupancy experiments.

Finally, models can be built for the C-terminal domain of mammalian CaM mutants that are associated with disease, again using the same basic structure of the model and only changing the rate parameters. A model with rate parameters that accurately simulate the calcium affinity and calcium dissociation rate of N54I CaM was created (Figure 7). N54I CaM has a calcium affinity slightly lower than wild type CaM, and a calcium dissociation rate that is slightly slower than wild type CaM. Since a decrease in $k_{off,2}$ increases the calcium affinity, a concurrent increase in $k_{off,1}$ was introduced to accurately simulate the experimentally determined affinity. A model with rate parameters that accurately simulate the calcium affinity and calcium dissociation rate of N98S CaM was also created (Figure 8). This model is more interesting than some of the others. When the same basic structure of the model is used, the simulated calcium titration curve is steeper than the experimental curve, indicative of excess cooperativity. Thus, the basic structure of the model must be changed, since simply changing the rate parameters to deplete the cooperativity (i.e., make $k_{on,1} \approx k_{on,2}$ and $k_{off,1} \approx k_{off,2}$) does not flatten the curve enough. When the basic structure is changed so that there is only one calcium ion binding to CaM, the steepness of the simulated titration curve matches the experimental curve. Here a modeling approach allows the conclusion to be reached that the N98S mutation hinders or abolishes calcium binding at one of the two binding sites in the C-terminal domain, a conclusion which can be further supported by noting that the 98^{th} residue of CaM is part of the loop of one of the helix-loop-helix calcium binding motifs. The relationship between the experimental data and the rate parameters in the one-calcium model is more direct – the experimentally determined calcium dissociation rate is still k_{off}, and since there is only one step

k_{on} can be calculated from the K_D using the formula $K_D = k_{off}/k_{on}$. While there is no experimental data for calcium association rates, if the one-calcium model is correct, the simulated calcium association rates should match any experimentally determined rates.

Wild Type and Other Calmodulin Variants – N Domain

Since these mathematical models only simulate individual domains of CaM, it is necessary to create a different model for the calcium binding properties of the N-terminal domain as well. The N-terminal domain shares the same general properties as the C-terminal domain, specifically, it binds two calcium ions as well and exhibits a similar level of cooperativity, so the same basic structure of the model can be used. Using this basic structure, a model with rate parameters that accurately simulate the calcium affinity and calcium dissociation rate of the N-terminal domain of wild-type CaM was created (Figure 9). Comparing this to the C-terminal domain, the calcium dissociation rate of the N-terminal domain is almost two orders of magnitude faster than the C-terminal domain, but the affinity is only ~2-fold lower. Thus, the high $k_{off,1}$ and $k_{off,2}$ values of the N-terminal domain are complemented with faster $k_{on,1}$ and $k_{on,2}$ values, with $k_{on,2}$ approaching the diffusion-controlled limit for a rate constant. That being said, there is no experimental data for calcium association rates, so the rate parameters presented here are only one of a large set of possible solutions to these models, and it is possible that these rate parameters would not accurately simulate calcium associate rate or transient occupancy experiments. To put a positive spin on this, simulations with this model can be used to predict experimental data, which has its own merits.

The same basic model structure can of course also be applied to the N-terminal domain of the different CaM variants. A model with rate parameters that accurately simulate the calcium binding properties of the N-terminal domain of sCaM1 was created (Figure 10). The calcium dissociation rate of sCaM1 is only slightly faster than the calcium dissociation rate of wild-type

CaM, and the affinity is only slightly higher than that of wild-type CaM. Due to the difference in calcium dissociation rate, $k_{off,2}$ is slightly higher for sCaM1, and interestingly, to accurately simulate the experimental calcium association data, $k_{on,1}$ must be slower than wild type. Both of these changes would decrease the affinity compared to wild type. So, to compensate for this and increases the affinity, $k_{off,1}$ is much slower than wild type, and is even slower than $k_{off,2}$ in this model, which is not consistent with the principles of cooperativity that must be incorporated into the model. However, it is important to note that there is calcium association rate data for the N-terminal domain of sCaM1, so there is higher confidence in the value for $k_{on,1}$ and $k_{on,2}$.

Finally, a model with rate parameters that effectively simulate the calcium binding properties of the N-terminal domain of sCaM4 was created (Figure 11). The calcium dissociation rate of sCaM4 is ~2-fold slower than wild-type CaM, and the calcium affinity is ~2.5-fold higher. Similar to sCaM1, the most interesting aspect of this model is the calcium association rates. In order to simulate these data, $k_{on,1}$ must be slower than $k_{on,1}$ for wild type, so $k_{off,1}$ must decrease to maintain the correct calcium affinity in the model. This leads to $k_{off,1}$ being slower to $k_{off,2}$, which again is not consistent with the principles of cooperativity that must be incorporated into the model. However, the presented parameters only accurately simulate half of the experimental data for calcium association. Given the presented basic structure of the model, there is no set of rate parameters that can accurately simulate all of the calcium association data. Perhaps a mathematical model for the N-terminal domain of sCaM4 requires additional complexity to accurately simulate all experimental data.

Symmetrical Parallel Model of Wild Type C Domain

As mentioned previously, the series model of calcium binding was chosen to simulate the experimental data based on its relative simplicity compared to the parallel model, and the fact that

it is actually possible to obtain a single solution to a model given a complete set of experimental data. However, creating parallel models that can accurately simulate the calcium binding data of CaM can be interesting as a thought experiment at the very least. There are two approaches to building a parallel model: it can be assumed that the two branches are symmetrical, meaning they have identical rate parameters, or the two branches can be asymmetrical, meaning they have different rate parameters. A symmetrical model is slightly simpler, so this approach was used first. The set of rate parameters within the parallel structure that create a symmetrical model that accurately simulates the calcium binding properties of the C-terminal domain of wild-type CaM are presented in Figure 12. As expected, the individual rate parameters are slower than their "corresponding" rate parameters in the series model, since there are two paths for complete calcium binding. But of course, a symmetrical model with this parameter set is not the only solution to the parallel model structure that can accurately simulate the calcium binding properties of CaM.

Asymmetrical Parallel Models of Wild Type C Domain

Since there are more parameters in the parameter set for a parallel model, there is a large number of solutions to the model given a complete array of experimental data. In addition to the symmetrical model described above, the model can be asymmetrical, and in a practically infinite number of ways. Two examples of sets of rate parameters within the parallel structure that create an asymmetrical model that accurately simulates the calcium binding properties of the C-terminal domain of wild-type CaM are presented in Figure 13. The first is "less" asymmetrical, with the affinity of the binding of the first calcium to site one approximately 1.5 times higher than the affinity for site 2. The second example is "more" asymmetrical, with the affinity of binding to site one approximately 45 times higher than the affinity for site 2. These two examples are only two possibilities for a solution to the model – there are an infinite number of parameter sets that can

accurately simulate all of the available experimental data. There are currently no known experiments that would allow for further characterization of the individual reactions seen in the parallel structure, so only a sample of potential solutions can be provided.

Slow Skeletal Troponin Model

As mentioned previously, a mathematical model of the activation of cardiac troponin has been created by a former graduate student in the Davis lab. This model works well for the cardiac isoforms of the troponin proteins, but it is important to understand the kinetics of the activation of troponin in other muscle types. Skeletal muscle disease is widely prevalent, and thus describing the calcium binding properties of skeletal muscle troponin is useful. For this project, slow skeletal troponin isoforms were chosen over the fast skeletal isoforms because in the long run, focusing on slow skeletal muscle fibers has more promise than focusing on fast skeletal muscle fibers. The different isoforms between these two muscle types lead to differences in their calcium binding properties – slow skeletal troponin has a higher calcium affinity when considering both the isolated troponin complex (~3-fold higher) and the thin filament (~1.2-fold higher) and slow skeletal troponin has a slower calcium dissociation rate when considering both the isolated troponin complex (~5-fold slower) and the thin filament (~1.3-fold slower). Given the relative complexity of the model, it was important to narrow down which parameters were candidates for manipulation. Since the TnC isoform is the same in cardiac and slow skeletal muscle, any rate parameters describing the kinetics of calcium binding to TnC were assumed to remain the same, as well as the dissociation of the TnC-TnI state back into its individual components. Thus, only $k_{on,2}$ and $k_{off,2}$ were subject to manipulation, and $k_{off,3}$ was as well because of the significant decrease in calcium dissociation rate.

The first step was to decrease $k_{off,3}$, since the $k_{off,3}$ value for the cardiac model was too high to work with the slow calcium dissociation rate of slow skeletal troponin. Next was to manipulate both $k_{on,2}$ and $k_{off,2}$ in the troponin complex to match the K_D and the calcium dissociation rate of the isolated slow skeletal troponin complex, maintaining the effective concentration of TnI at the same value as it was for the cardiac troponin complex, since it is assumed that the troponin complex has saturating TnI. Since there is no experimental data for calcium association rate, there were multiple sets of rate parameters that would match the other experimental data (however, not an infinite number). The number of viable sets was cut down to two options: (1) $k_{on,2} = 2.955 \times 10^7$ M^{-1} s^{-1} and $k_{off,2} = 110$ s^{-1}, since $k_{off,2} = 110$ s^{-1} in the cardiac model and $k_{off,4} = 110$ s^{-1}; and (2) $k_{on,2} = 2.313 \times 10^7$ M^{-1} s^{-1} and $k_{off,2} = 85$ s^{-1}, since the calcium dissociation rate for the slow skeletal thin filament is 85 s^{-1}, and the same logic was used when determining $k_{off,2}$ for the cardiac model (i.e.. $k_{off,2}$ is equal to the calcium dissociation rate for the thin filament). These two options were translated to the thin filament, where an extra parameter played a role in determining calcium affinity and dissociation rate: the effective concentration of TnI. Using the same two sets of $k_{on,2}$ and $k_{off,2}$, an effective concentration of TnI was calculated for each set that would accurately simulate the calcium affinity and dissociation rate. Thus, there are essentially two viable models that can describe the calcium binding properties of slow skeletal troponin (Table 1).

As anticipated, calcium association data is needed to further determine which of the two models is accurate. Interestingly, if the two models are used to simulate transient occupancy in the thin filament system, they give two significantly different values, one of which is much closer to the expected/simulated transient occupancy of the cardiac thin filament system (Figure 14). Therefore, it appears that a simple transient occupancy experiment with slow skeletal thin filament

would be sufficient to identify which of the two potential models describes the calcium binding properties more accurately.

Cardiac Troponin Drug Model

In addition to translating the previously made cardiac model into a slow skeletal model, it can be extended to include reactions and species that correspond to a small molecule drug binding to the closed state of TnC-Ca (Figure 15). A drug model can be useful to evaluate potential small molecule therapeutics and determine whether certain theories about the small molecule binding events are viable. In this case, the model was used to evaluate a single small molecule and its effects on the calcium binding properties of troponin. The small molecule was selected for evaluation based on computational models and high throughput screening of databases, and its interactions with troponin had previously been characterized through biochemical and structural (NMR) studies. To accurately describe its interactions, a clarification must be added to the model. The state TnC-Ca^{2+} in the original troponin model is a simplification – this state exists in either a "closed" or an "open" state, with an exchange rate of 30000 per second, and the closed state is 5% of the population in saturating calcium conditions (Cai et al., 2016). Returning to the small molecule's interactions, based on the NMR studies, it can bind troponin in two states: (1) TnC-Ca^{2+}-Closed and (2) TnC-Ca^{2+}-TnI (Cai et al., 2016). It is because of this second interaction that it was chosen for further studies, binding and subsequently stabilizing the active form of troponin should theoretically increase the calcium sensitivity of troponin, which is what is desired for a potential therapeutic. The small molecule binds the TnC-Ca^{2+}-Closed state with an exchange rate of 1000 per second and a dissociation rate of 0.3 per second, giving an assumed K_D of 300 μM, and it binds the TnC-Ca^{2+}-TnI state with a k_{off} of 1000 per second and a K_D of 12 μM, giving an assumed k_{on} of 8.33×10^7 M^{-1} s^{-1} (Cai et al., 2016). When in the subsequent TnC-Ca^{2+}-TnI-Drug

state, calcium dissociates at a rate of 10 per second, and it was assumed that the TnC-TnI-Drug state falls apart quickly at the same rate as TnC-TnI, 110 per second.

Once the small molecule had been added to the model, several different simulations were performed to try and validate experimental data and hypotheses of how the small molecule conferred its effects. There were a couple important takeaways from these simulations. Primarily, it was determined that in addition to sensitizing the thin filament to calcium, addition of the small molecule increased the maximum activation of troponin at saturating calcium conditions (Figure 16). In the thin filament, where this effect is more pronounced, the effective concentration of TnI is only ~10-fold higher than the K_D of the TnI binding reaction ($[TnI]_{eff}$ = 12.3 μM, $K_{D,2}$ = 1.22 μM). Thus, without the small molecule, at saturating calcium conditions only ~80-90% of troponin is fully activated. When the small molecule is added, it shifts the equilibrium towards the TnC-Ca^{2+}-TnI state and increases the maximum activation of the thin filament (i.e., the percentage of troponin that is fully active at saturating calcium). In addition to examining the sensitizing effects of the small molecule, it was important to look at the inhibitory effects as well, since it binds the closed state of TnC-Ca^{2+}. It was hypothesized that the small molecule would have a significant inhibitory effect but based on the parameters assumed in the mathematical model, this was not the case. The population of troponin that ends up in the deactivated state in equilibrium is practically insignificant with the given parameters (Figure 17), and the affinity of the inhibitory reaction (small molecule binding the TnC-Ca^{2+}-Closed state) has to increase significantly to give a relevant population of deactivated species at physiologically relevant small molecule concentrations (Table 2).

Chapter 4 - Discussion

As demonstrated, mathematical modeling has several merits and can be incredibly useful to describe and gain information about given biochemical systems. In this project, the biochemical systems modeled were two calcium binding protein systems – the two individual domains of calmodulin, and different isoforms and applications of the proteins in the troponin complex. For the different CaM variants, some models have been experimentally verified with a complete set of calcium binding data (i.e., steady state affinity, dissociation rate, and association rate), and some are lacking one set of data to verify the accuracy of the model. While it may seem that unverified models may not be useful, there are some new avenues that open up. These models can be used to predict experiments, which can either verify the model completely or indicate the need for change, much like the situation with the slow skeletal troponin model. Even then, with completely verified models, there is still predictive value. For example, a potential experiment may be tested and simulated with the model prior to physically performing it to see if there is merit to running the experiment in the first place. Returning to the slow skeletal example, a thin filament transient occupancy experiment was not the only one considered to potentially distinguish between the two models, so the models were used to determine that it was the only viable option as opposed to running experiments that would end up yielding no new results.

It is this concept of predictive value that can be seen as the most useful aspect of models. In addition to considering what experiments to do in the future, the models can be used to design new experiments to answer new questions that may be too difficult to theorize about. In addition to noting how these models can simulate these calcium binding protein systems and aid in future experiments, it is also interesting to consider some of the different potential structures of the CaM model and to investigate the mathematical relationships between the parameters of the model.

Series vs. Parallel Models

As mentioned previously, there were two options to consider in terms of a structure for the model of calcium binding to CaM. Between the series and the parallel model, the series model was chosen as it was simpler with fewer parameters. However, if the point of the model is to represent calcium binding to individual domains as accurately as possible, a parallel model would be the better option. In essence, a series model is simply a parallel model with the rate parameters of one branch all equal to 0, which would indicate that the first calcium ion that binds CaM will always bind at one of the two sites, and the other binding site would always bind the second calcium ion that arrives. Unsurprisingly, this assumption does not have much validity to it. There is no structural or biochemical reason that the first calcium ion coming in would always head to one binding site over the other. There is no occlusion of either site in the apo state, and each calcium binding loop motif is typical, in the sense that they both have amino acids that can effectively coordinate calcium ions without the aid of other structural motifs. Thus, it is clear that the more accurate description of calcium binding to CaM is given by a parallel model, even though it is more complex than a series model and there is not an effective way to determine a single solution to the parameter set. But even within the frame of a parallel model, there are multiple ways that these reactions can be described.

Symmetrical vs. Asymmetrical Models

Given a parallel model, there are several parameter sets that can accurately simulate the calcium binding properties of the individual domains of CaM. While there are practically an infinite number of solutions to the parallel model, some can be narrowed down by thinking conceptually and by examining the mathematical relationships between the different parameters. The first concept that can be examined is whether the parallel model should be symmetrical or

asymmetrical. This project has shown that both options are viable for simulating the calcium binding properties required of the model, but it is likely that one of the options makes more sense conceptually. If a symmetrical model is considered, an implication of this is that both pathways to full calcium binding are equal – each calcium binding loop would have the same affinity. This assumption is likely to be invalid. The sequences of the two calcium binding loops are not identical; thus, it should not be too far-fetched to assume that the affinity of the individual loops for calcium are also not identical. Experimental validation for this assumption has been difficult to come by given the fleeting nature of single calcium-bound CaM, but efforts have been made with mutational and thermodynamic studies. Following mutagenesis and typical spectrofluorimetric titrations, it has been demonstrated that the affinity of the two calcium binding sites in the N-terminal domain of CaM differ by approximately 1.5-fold (Beccia et al., 2015). While it is possible to do a similar analysis with the C-terminal domain using the N98S mutation and another mutation in the other calcium binding loop in the C-terminal domain (ex. D130V, D132V, or Q136P) given that the N98S mutation appears to knock out a binding site, it is unclear if this analysis would work because some of the other mutations can have effects that are more drastic than simply knocking out one binding site. Thus, it is unclear if the magnitude of this difference translates to the C-terminal domain as well. Nevertheless, these results act as a proof of concept at the very least – the two calcium binding loops in either domain of CaM do not have the same affinity. Therefore, it is likely that an asymmetric parallel model is a more accurate description of calcium binding to CaM than a symmetric parallel model.

So, through conceptualization of the calcium binding process the most accurate model has been narrowed down to an asymmetric parallel model. But, consistent with the theme of a parallel model, there are many different parameter sets than can accurately simulate the calcium binding

properties of CaM in an asymmetrical manner. There are several questions that can be asked: How asymmetrical is the model/what is the magnitude of the difference between branches? Which group of parameters should be asymmetrical, or should all of them be asymmetrical? Clearly, there needs to be some sort of experimental validation when creating a parallel model. Unfortunately, at this moment in time no such experiments have been performed, and it is difficult to even think of a method other than potentially some sort of NMR experiment that can be used to uncover information about these binding events at such a microscopic level.

Mathematical Considerations

To further characterize the asymmetrical parallel model structure, several different asymmetrical models were created (see Figure 13 for two examples). Throughout this process and upon analyzing the results, some notable observations were made. Initially, to create an asymmetrical model, changes were made to the symmetrical model such that $k_{1,off}$ was decreased to increase the affinity for calcium within the first branch and $k_{1,2,on}$ was decreased to decrease the affinity for calcium within the second branch. After the initial changes, small tweaks were made in these parameters to ensure the affinity and calcium dissociation rate matched the experimental data, and $k_{1,on}$ was changed slightly to ensure the calcium association rates could be accurately simulated. This approach could be done to make the model as asymmetrical as possible. However, it was noted that if this approach was taken, the simulated calcium titration curves began to become less steep – the models were losing cooperativity. Because of this, a different approach had to be taken to ensure that the cooperativity aspect of the calcium binding is accurate.

This is where mathematical relationships between the rate parameters in the parallel model can be useful. These mathematical relationships can be derived from the mathematical relationships that are true for any calcium binding system that has two binding sites. To summarize

these relationships, there are the microscopic association constants $K_1 = k_{1,on}/k_{1,off}$, $K_2 = k_{2,on}/k_{2,off}$, $K_{2,1} = k_{2,1,on}/k_{2,1,off}$, and $K_{1,2} = k_{1,2,on}/k_{1,2,off}$, and these microscopic association constants are related to the macroscopic association constants with the following equations (Linse & Forsén, 1995):

$$K_I = \frac{[Ca_1CaM]}{[CaM][Ca]} = K_1 + K_2$$

$$K_{II} = \frac{[Ca_2CaM]}{[Ca_1CaM][Ca]} = \frac{K_1 * K_{2,1}}{K_1 + K_2}$$

Using the first relationship, an asymmetrical model can be built by setting the sum of K_1 and K_2 of the symmetrical model equal to the sum of K_1 and K_2 of the new asymmetrical model and solving for a new $k_{1,on}$ after assigning a new $k_{2,on}$. No matter what value is assigned to $k_{2,on}$, using this relationship to solve for $k_{1,on}$ will give an asymmetrical model that simulates all calcium binding properties, including cooperativity. Essentially, as long as $k_{1,on} + k_{2,on}$ is equal to some constant, any set of values for those two parameters will work for the model. The same sort of relationship is seen with $k_{1,2,off}$ and $k_{2,1,off}$. As long as those two parameters sum to a constant (which in this case is the experimental calcium dissociation rate), the model will accurately simulate all of the calcium binding properties of CaM.

Interestingly, a similar relationship is not seen between $k_{1,2,on}$ and $k_{2,1,on}$ or between $k_{1,off}$ and $k_{2,off}$. In fact, it has proven to be difficult to determine what relationship there is. Using a derivation of the second equation from above leads to a model with slightly altered calcium affinity. Other potential relationships such as $k_{1,off} * k_{2,off}$ equaling a constant lead to the same outcome – slightly altered affinity. While it is possible to assign a new value to a parameter such as $k_{1,off}$ and have a new value of $k_{2,off}$ that will maintain the same calcium affinity as well as the

other calcium binding properties, it is still unclear what the mathematical relationship between those two parameters is.

Slow Skeletal Troponin Insights

Although an exact model for the activation of slow skeletal troponin was not able to be created, there is still information that can be gathered based on the two models that are proposed. Primarily, it is important to note that $k_{off,3}$ was significantly lower in the slow skeletal model than in the cardiac model (by ~6-fold). This implies that the slow skeletal isoform of TnI allows for TnC to hold on to its calcium ion much better than the cardiac isoform of TnI. Another potential interpretation could be that the initial model for cardiac troponin had a value too high for $k_{off,3}$, because the argument could be made that since the reaction that $k_{off,3}$ describes is between TnC and calcium, $k_{off,3}$ should be unchanged between the two models since TnC is unchanged between the two models. However, this would mean that in order to match the experimental calcium dissociation rate for the cardiac troponin complex, the TnI effective concentration would be significantly lower than in the slow skeletal troponin complex, which is far-fetched and would throw a wrench in the process that was used to make the slow skeletal models to begin with, since it was assumed that the TnI effective concentration was saturating in the troponin complex and hence the same in the cardiac and slow skeletal isoforms.

The other important takeaways from the potential slow skeletal models center around the TnI binding step (step 2). Notably, in both potential models, $k_{on,2}$ is lower than $k_{on,2}$ in the cardiac model. Because the magnitude of the difference between the cardiac $k_{on,2}$ and the slow skeletal values for $k_{on,2}$ is greater than the potential difference in $k_{off,2}$, the slow skeletal models indicate that the affinity of slow skeletal TnI for TnC is lower than the affinity of cardiac TnI for TnC. But this is not the only point of interest. It is not unsurprising to see that it is likely that the effective

concentration of TnI is different in the cardiac thin filament and the slow skeletal thin filament, but it is interesting to see that it could be higher or lower, depending on the model. If $k_{off,2}$ is the same as in the cardiac model, then the effective concentration of TnI is higher in the slow skeletal thin filament than in the cardiac thin filament. But if $k_{off,2}$ is slower than the cardiac model, then the effective concentration of TnI is lower in the slow skeletal thin filament than in the cardiac thin filament. Clearly, even before a model has been concretely determined for the activation of slow skeletal troponin, there are many conclusions that can be drawn from the process of creating a model that can be of relevance in future studies.

Troponin Drug Model Takeaways

For the model of the activation of cardiac troponin that includes a small molecule, the model was used to show how the small molecule can sensitize troponin to calcium by both lowering the K_D and increasing the maximal activation. It was also used to demonstrate how the inhibitory effects of the small molecule were essentially negligible under physiologically relevant conditions. This result ends up being significant because it disproves the proposed mechanism of the small molecule binding. The mechanism that was initially proposed included the idea that there would be a significant inhibitory effect by the small molecule, but with the given rate parameters from the characterization of the small molecule the model disproves this theory. This is yet another way that the use of models can aid in understanding a biochemical system without performing wet lab experiments that would be costly and time consuming. The model allowed for additional characterization of the small molecule's interaction with troponin and vetting of the proposed mechanism, furthering the understanding of how it would act in a complex physiological system.

Conclusion & Future Directions

In this thesis project it has been demonstrated that mathematical models are both a valid approach to understanding and describing biochemical systems and useful to aid in data interpretation in a more detailed fashion. These models show that systems that can seem to be simple can actually be quite complex and that a lot of detail must go into their descriptions. They can demonstrate the impact of a mutation on a protein, they can allow for insight into microscopic binding events that differ between proteins, and they can predict and explain experimental data with precision. But there are of course other applications of the models presented in this thesis and expansions that can be made.

For the CaM models, in order to make them more physiologically relevant, they need to include more than only one domain of CaM and calcium. Next steps to this project include combining models for the C-terminal domain and the N-terminal domain into one model and adding one or more target proteins to the model. As discussed earlier, mathematical modeling of CaM and its target proteins can lead to important discoveries, even if the models are not as detailed as they could be. Creating a model with detailed calcium binding parameters that can answer physiologically relevant questions would allow for several important insights to be gained.

For the troponin models, the next step would be to experimentally validate which of the two slow skeletal models is more accurate by performing the transient occupancy experiments described previously and comparing the results from the slow skeletal and cardiac thin filaments. These experiments can be done with relative ease and would have been performed with more time. Once this has been done, the validated model can be used to answer several questions concerning slow skeletal troponin, both in isolation and compared to the cardiac isoform. For example, it can

be used to evaluate whether the conclusions based on TnI effective concentration in the cardiac model are valid in slow skeletal muscle as well.

In sum, this thesis has shown that mathematical modeling is a useful tool for describing and understanding biochemical systems, and that this approach shows promise for answering future questions. Many questions can be asked and answered using models that can both supplement other experimental approaches and be specific to modeling. Due to its relative simplicity and powerful abilities, a mathematical model is an approach that should be considered by researchers to aid in their quest for the discovery of new knowledge.

References

Al-Quraan, N., & Singh, N. K. (2008). *Role of Arabidopsis thaliana Calmodulin isoforms in tolerance to Abiotic Stress* (Doctoral dissertation, Auburn University, 2008). Ann Arbor, MI: ProQuest Dissertations Publishing.

Alves, M. L., Dias, F. A., Gaffin, R. D., Simon, J. N., Montminy, E. M., Biesiadecki, B. J., . . . Wolska, B. M. (2014). Desensitization of myofilaments to ca 2+ as a therapeutic target for hypertrophic cardiomyopathy with mutations in thin filament proteins. *Circulation: Cardiovascular Genetics, 7*(2), 132-143. doi:10.1161/circgenetics.113.000324

Beccia, M. R., Sauge-Merle, S., Lemaire, D., Brémond, N., Pardoux, R., Blangy, S., . . . Berthomieu, C. (2015). Thermodynamics of Calcium binding to the Calmodulin N-terminal domain to evaluate site-specific affinity constants and cooperativity. *JBIC Journal of Biological Inorganic Chemistry, 20*(5), 905-919. doi:10.1007/s00775-015-1275-1

Bogdanov, V., Tikunova, S., Kou, Y., Fadell, N., Evans, J., Tirone, A., . . . Davis, J. P. (2020). In silico engineering of Calmodulin to bind the Cardiac Ryanodine receptor with high affinity. *Biophysical Journal, 118*(3). doi:10.1016/j.bpj.2019.11.3215

Cai, F., Li, M. X., Pineda-Sanabria, S. E., Gelozia, S., Lindert, S., West, F., . . . Hwang, P. M. (2016). Structures reveal details of small molecule binding to cardiac troponin. *Journal of Molecular and Cellular Cardiology, 101*, 134-144. doi:10.1016/j.yjmcc.2016.10.016

Chin, D., & Means, A. R. (2000). Calmodulin: A Prototypical Calcium Sensor. *Trends in Cell Biology, 10*(8), 322-328. doi:10.1016/s0962-8924(00)01800-6

Cho, M. J., Vaghy, P. L., Kondo, R., Lee, S. H., Davis, J. P., Rehl, R., . . . Johnson, J. D. (1998). Reciprocal regulation of Mammalian nitric oxide synthase and Calcineurin by Plant Calmodulin Isoforms. *Biochemistry, 37*(45), 15593-15597. doi:10.1021/bi981497g

Davis, J. P., Norman, C., Kobayashi, T., Solaro, R. J., Swartz, D. R., & Tikunova, S. B. (2007b). Effects of thin and thick Filament proteins on Calcium binding and exchange with Cardiac Troponin C. *Biophysical Journal, 92*(9), 3195-3206. doi:10.1529/biophysj.106.095406

Davis, J. P., Shettigar, V., Tikunova, S. B., Little, S. C., Liu, B., Siddiqui, J. K., . . . Walton, S. D. (2016). Designing proteins to combat disease: Cardiac troponin C as an example. *Archives of Biochemistry and Biophysics, 601*, 4-10. doi:10.1016/j.abb.2016.02.007

Davis, J. P., & Tikunova, S. B. (2008). Ca2+ exchange with troponin c and cardiac muscle dynamics. *Cardiovascular Research, 77*(4), 619-626. doi:10.1093/cvr/cvm098

Davis, J., Wen, H., Edwards, T., & Metzger, J. M. (2007a). Thin filament disinhibition by restrictive cardiomyopathy mutant r193h troponin I induces ca 2+ -independent mechanical

tone and acute myocyte remodeling. *Circulation Research, 100*(10), 1494-1502. doi:10.1161/01.res.0000268412.34364.50

Du, C., Morimoto, S., Nishii, K., Minakami, R., Ohta, M., Tadano, N., . . . Sasaguri, T. (2007). Knock-in mouse model of dilated cardiomyopathy caused by troponin mutation. *Circulation Research, 101*(2), 185-194. doi:10.1161/circresaha.106.146670

Fischer, R., Koller, M., Flura, M., Mathews, S., Strehler-Page, M. A., Krebs, J., . . . Strehler, E. E. (1988). Multiple divergent mRNAs code for a single human calmodulin. *Journal of Biological Chemistry, 263*(32), 17055-17062. doi:10.1016/s0021-9258(18)37497-0

Gomes, A. V., Venkatraman, G., Davis, J. P., Tikunova, S. B., Engel, P., Solaro, R. J., & Potter, J. D. (2004). Cardiac troponin T Isoforms affect the Ca2+ sensitivity of force development in the presence of Slow Skeletal Troponin I. *Journal of Biological Chemistry, 279*(48), 49579-49587. doi:10.1074/jbc.m407340200

Jin, J. P. (2014). Troponin: Regulator of muscle contraction. New York: Nova Biomedical.

Johnson, J. D., Snyder, C., Walsh, M., & Flynn, M. (1996). Effects of Myosin light Chain kinase and Peptides on Ca2+ exchange with the N- and C-Terminal Ca2+ binding sites of Calmodulin. *Journal of Biological Chemistry, 271*(2), 761-767. doi:10.1074/jbc.271.2.761

Kondo, R., Tikunova, S. B., Cho, M. J., & Johnson, J. D. (1999). A point mutation in a plant calmodulin is responsible for its inhibition of nitric-oxide synthase. *Journal of Biological Chemistry, 274*(51), 36213-36218. doi:10.1074/jbc.274.51.36213

Layland, J., Solaro, R., & Shah, A. (2005). Regulation of cardiac contractile function by troponin I phosphorylation. *Cardiovascular Research, 66*(1), 12-21. doi:10.1016/j.cardiores.2004.12.022

Lee, S. H., Johnson, J. D., Walsh, M. P., Van Lierop, J. E., Sutherland, C., Xu, A., . . . Cho, M. J. (2000). Differential regulation of Ca2+/calmodulin-dependent enzymes by plant calmodulin isoforms and free Ca2+ concentration. *Biochemical Journal, 350*(1), 299-306. doi:10.1042/bj3500299

Lee, S. H., Kim, J. C., Lee, M. S., Heo, W. D., Seo, H. Y., Yoon, H. W., . . . Cho, M. J. (1995). Identification of a novel divergent calmodulin isoform from soybean which has differential ability to activate calmodulin-dependent enzymes. *Journal of Biological Chemistry, 270*(37), 21806-21812. doi:10.1074/jbc.270.37.21806

Li, A. Y., Stevens, C. M., Liang, B., Rayani, K., Little, S., Davis, J., & Tibbits, G. F. (2013). Familial hypertrophic Cardiomyopathy Related Cardiac Troponin C L29q Mutation Alters Length-Dependent activation and Functional effects of phosphomimetic troponin I*. *PLoS ONE, 8*(11). doi:10.1371/journal.pone.0079363

Linse, S., & Forsén, S. (1995). Determinants that govern high-affinity calcium binding. *Advances in Second Messenger and Phosphoprotein Research, 30*, 89-151. doi:10.1016/s1040-7952(05)80005-9

Little, S. C. (2012). *The role of troponin c in the heart* (Doctoral dissertation, Ohio State University, 2012). Columbus, OH: OhioLink.

Liu, B., Lee, R. S., Biesiadecki, B. J., Tikunova, S. B., & Davis, J. P. (2012a). Engineered troponin c constructs correct disease-related cardiac myofilament calcium sensitivity. *Journal of Biological Chemistry, 287*(24), 20027-20036. doi:10.1074/jbc.m111.334953

Liu, B., Lopez, J. J., Biesiadecki, B. J., & Davis, J. P. (2014). Protein kinase c phosphomimetics alter thin filament ca2+ binding properties. *PLoS ONE, 9*(1). doi:10.1371/journal.pone.0086279

Liu, B., Tikunova, S. B., Kline, K. P., Siddiqui, J. K., & Davis, J. P. (2012b). Disease-related cardiac troponins alter thin filament ca2+ association and dissociation rates. *PLoS ONE, 7*(6). doi:10.1371/journal.pone.0038259

McCormack, E., & Braam, J. (2003). Calmodulins and related potential calcium sensors of Arabidopsis. *New Phytologist, 159*(3), 585-598. doi:10.1046/j.1469-8137.2003.00845.x

McCormack, E., Tsai, Y., & Braam, J. (2005). Handling calcium signaling: Arabidopsis cams and CMLs. *Trends in Plant Science, 10*(8), 383-389. doi:10.1016/j.tplants.2005.07.001

Ohki, S., Ikura, M., & Zhang, M. (1997). Identification of Mg2+-binding sites and the role of Mg2+on target recognition by calmodulin. *Biochemistry, 36*(14), 4309-4316. doi:10.1021/bi962759m

Park, H. C., Kim, M. L., Kang, Y. H., Jeong, J. C., Cheong, M. S., Choi, W., . . . Yun, D. (2009). Functional analysis of the stress-inducible soybean calmodulin isoform-4 (gmcam-4) promoter in transgenic tobacco plants. *Molecules and Cells, 27*(4), 475-480. doi:10.1007/s10059-009-0063-6

Pepke, S., Kinzer-Ursem, T., Mihalas, S., & Kennedy, M. B. (2010). A dynamic model of interactions of Ca2+, calmodulin, and Catalytic subunits Of Ca2+/Calmodulin-Dependent protein Kinase II. *PLoS Computational Biology, 6*(2). doi:10.1371/journal.pcbi.1000675

Pharris, M. C., Patel, N. M., & Kinzer-Ursem, T. L. (2018). Competitive tuning among ca2+/calmodulin-dependent proteins: Analysis of in silico model robustness and parameter variability. *Cellular and Molecular Bioengineering, 11*(5), 353-365. doi:10.1007/s12195-018-0549-4

Robertson, S., & Potter, J. D. (1984). The regulation of Free Ca2+ ion concentration by Metal Chelators. *Myocardial Biology,* 63-75. doi:10.1007/978-1-4684-4778-1_6

Romano, D. R., Pharris, M. C., Patel, N. M., & Kinzer-Ursem, T. L. (2017). Competitive tuning: Competition's role in setting the frequency-dependence of Ca2+-dependent proteins. *PLOS Computational Biology, 13*(11). doi:10.1371/journal.pcbi.1005820

Shettigar, V., Zhang, B., Little, S. C., Salhi, H. E., Hansen, B. J., Li, N., . . . Davis, J. P. (2016). Rationally engineered Troponin C modulates in vivo cardiac function and performance in health and disease. *Nature Communications, 7*(1). doi:10.1038/ncomms10794

Siddiqui, J. K., Tikunova, S. B., Walton, S. D., Liu, B., Meyer, M., De Tombe, P. P., . . . Davis, J. P. (2016). Myofilament calcium Sensitivity: Consequences of the effective concentration of Troponin I. *Frontiers in Physiology, 7*. doi:10.3389/fphys.2016.00632

Slavov, N., Carey, J., & Linse, S. (2013). Calmodulin transduces Ca2+ Oscillations into Differential regulation of its target proteins. *ACS Chemical Neuroscience, 4*(4), 601-612. doi:10.1021/cn300218d

Tikunova, S. B., Black, D. J., Johnson, J. D., & Davis, J. P. (2001). Modifying Mg2+ Binding and exchange with the N-terminal of calmodulin. *Biochemistry, 40*(11), 3348-3353. doi:10.1021/bi0021333

Tikunova, S. B., & Davis, J. P. (2004). Designing calcium-sensitizing mutations in the regulatory domain of cardiac troponin c. *Journal of Biological Chemistry, 279*(34), 35341-35352. doi:10.1074/jbc.m405413200

Vogel, H. J. (1994). Calmodulin: A Versatile calcium Mediator protein. *Biochemistry and Cell Biology, 72*(9-10), 357-376. doi:10.1139/o94-049

Walton, S. D., Chakravarthy, H., Shettigar, V., O'Neil, A. J., Siddiqui, J. K., Jones, B. R., . . . Davis, J. P. (2017). Divergent soybean calmodulins respond similarly to calcium transients: Insight into differential target regulation. *Frontiers in Plant Science, 08*. doi:10.3389/fpls.2017.00208

Waters, B. M. (2011). Moving magnesium in plant cells. *New Phytologist, 190*(3), 510-513. doi:10.1111/j.1469-8137.2011.03724.x

Figures and Tables

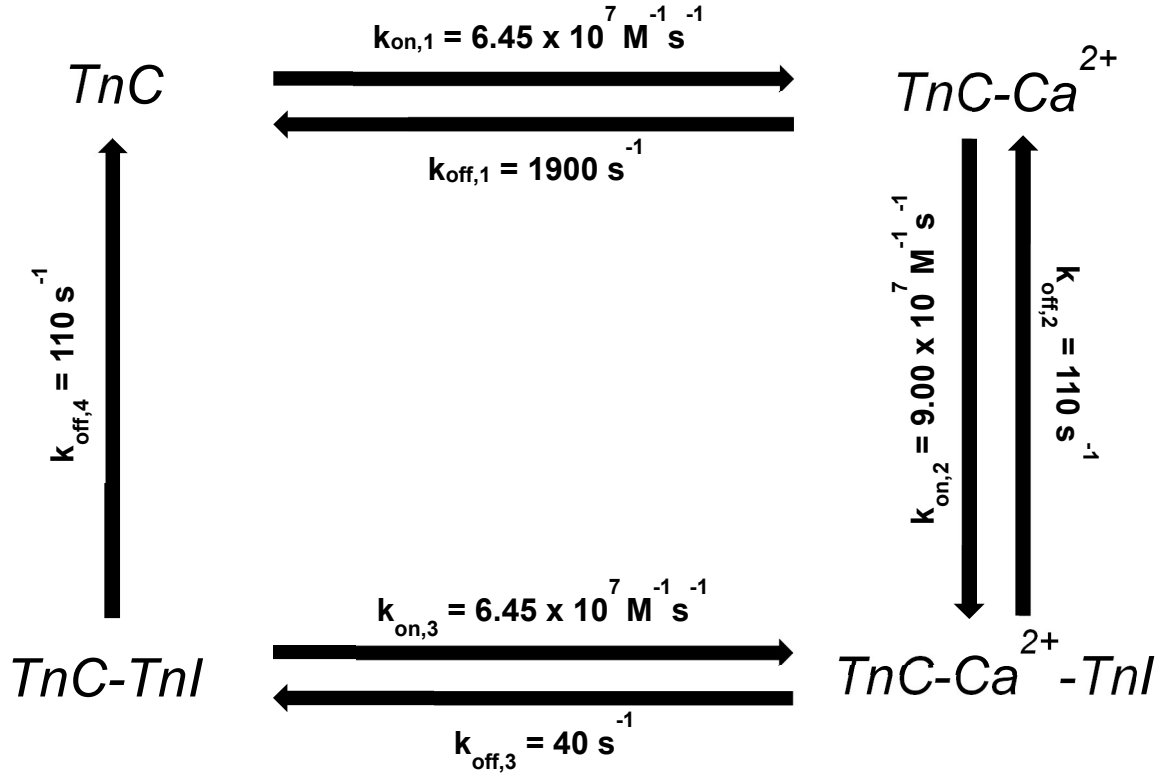
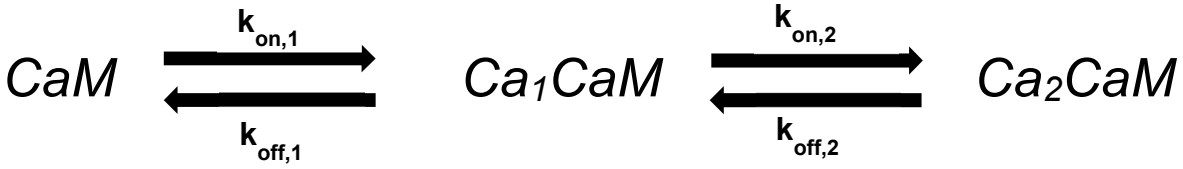


Figure 1. Schematic of a previously made mathematical model for the activation of cardiac troponin.

Series:



Parallel:

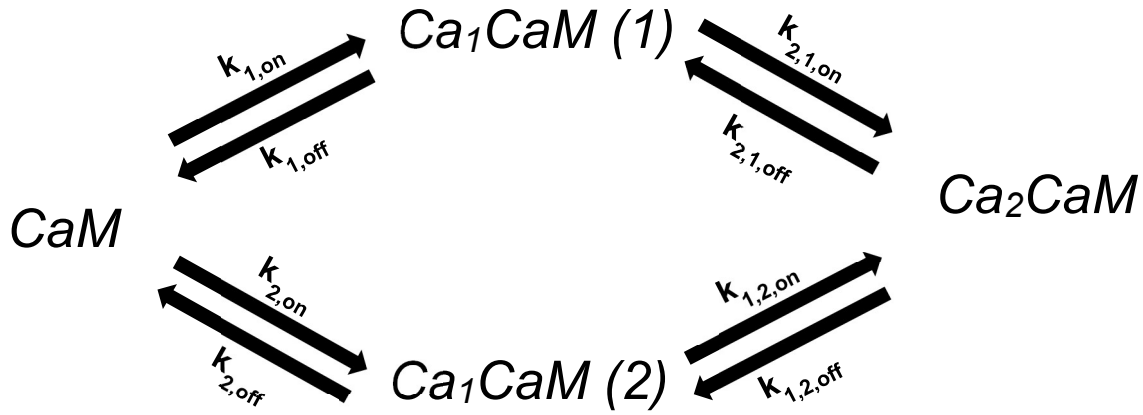


Figure 2. Schematics of a series and parallel model to describe calcium binding to CaM.

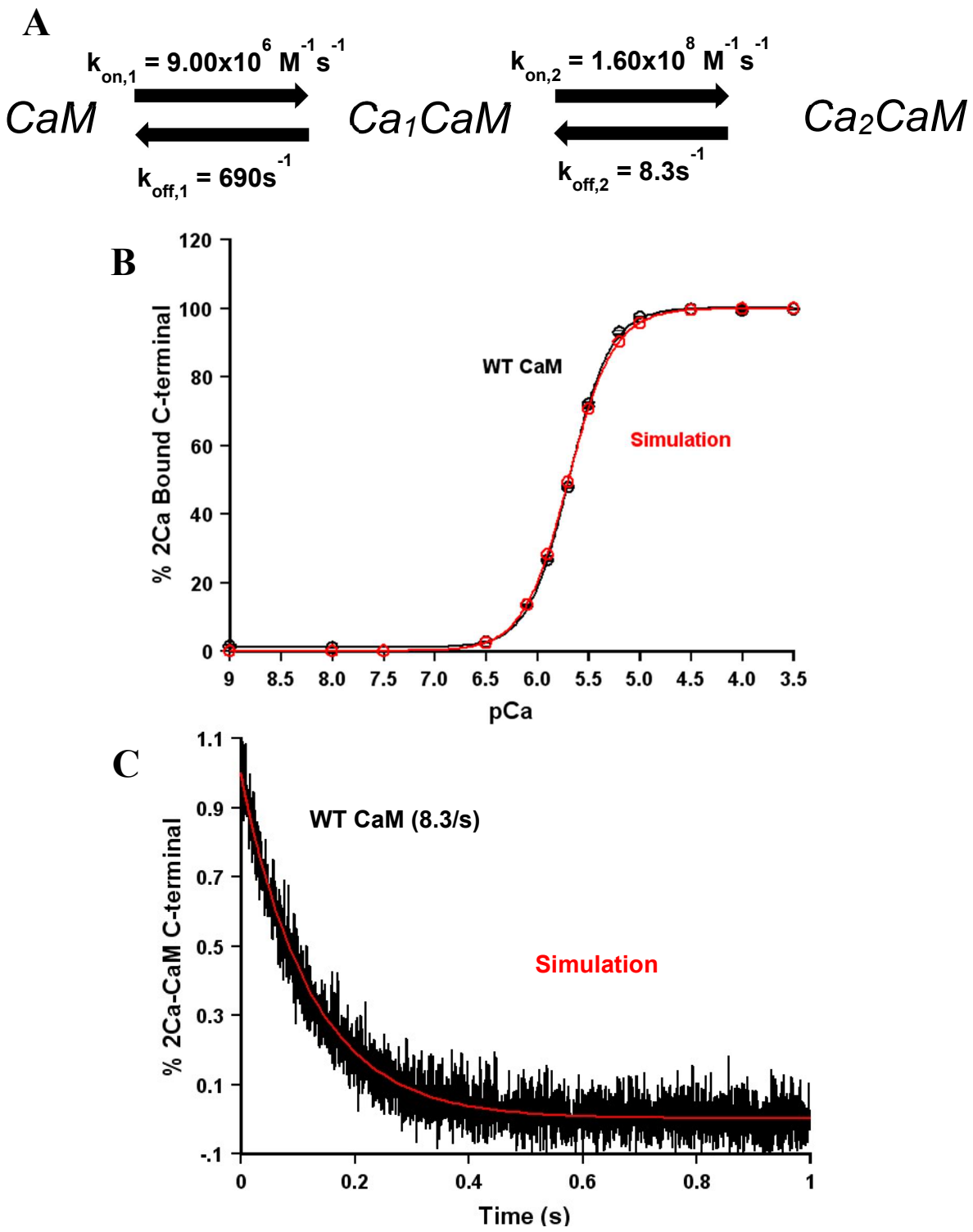


Figure 3. Summary of the model for the C-terminal domain of wild type CaM. (a) Schematic of the model with rate parameters, (b) Overlay of real steady-state affinity data with a simulated curve, and (c) Overlay of real calcium dissociation data with a simulated curve.

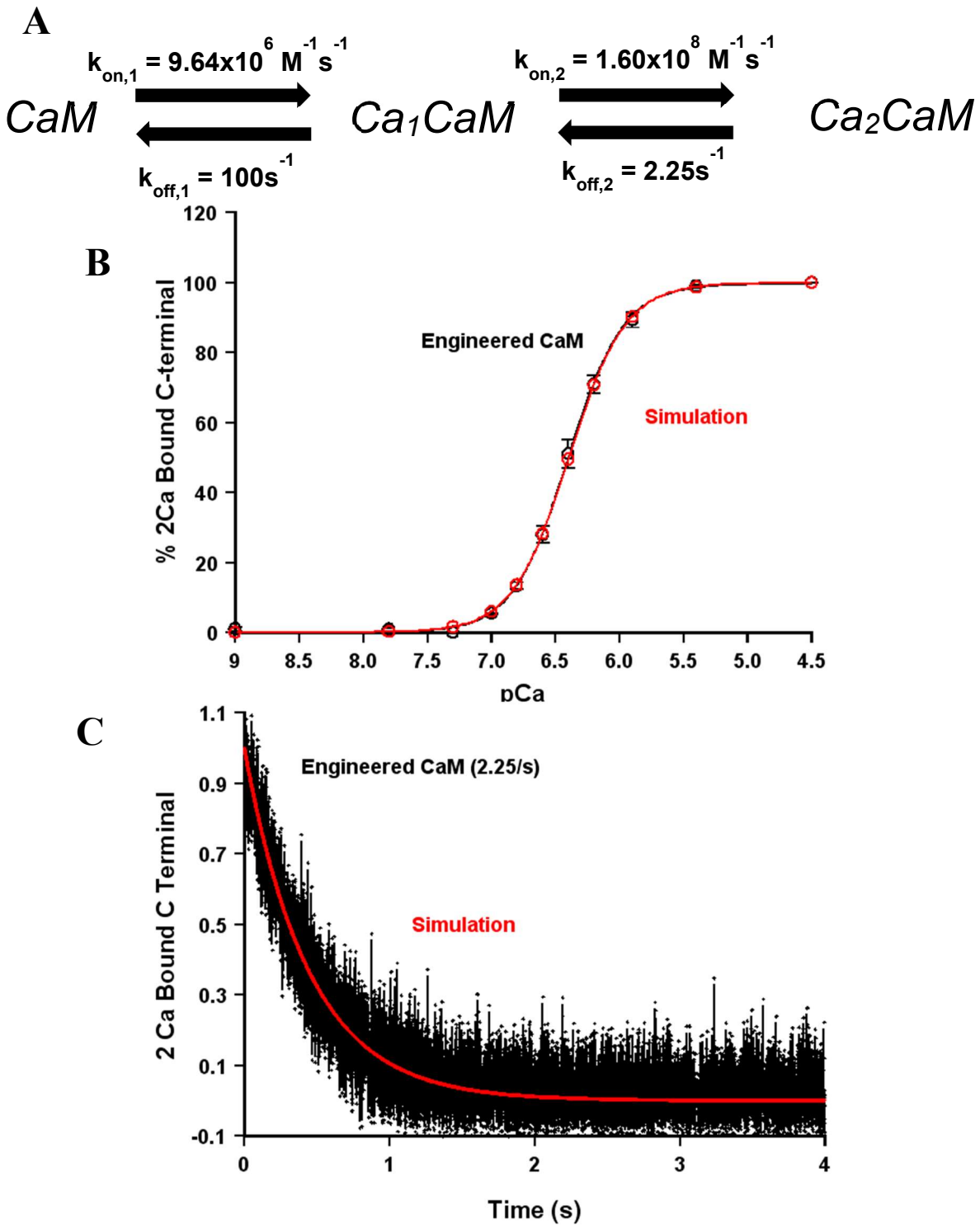


Figure 4. Summary of the model for the C-terminal domain of the engineered CaM. (a) Schematic of the model with rate parameters, (b) Overlay of real steady-state affinity data with a simulated curve, and (c) Overlay of real calcium dissociation data with a simulated curve.

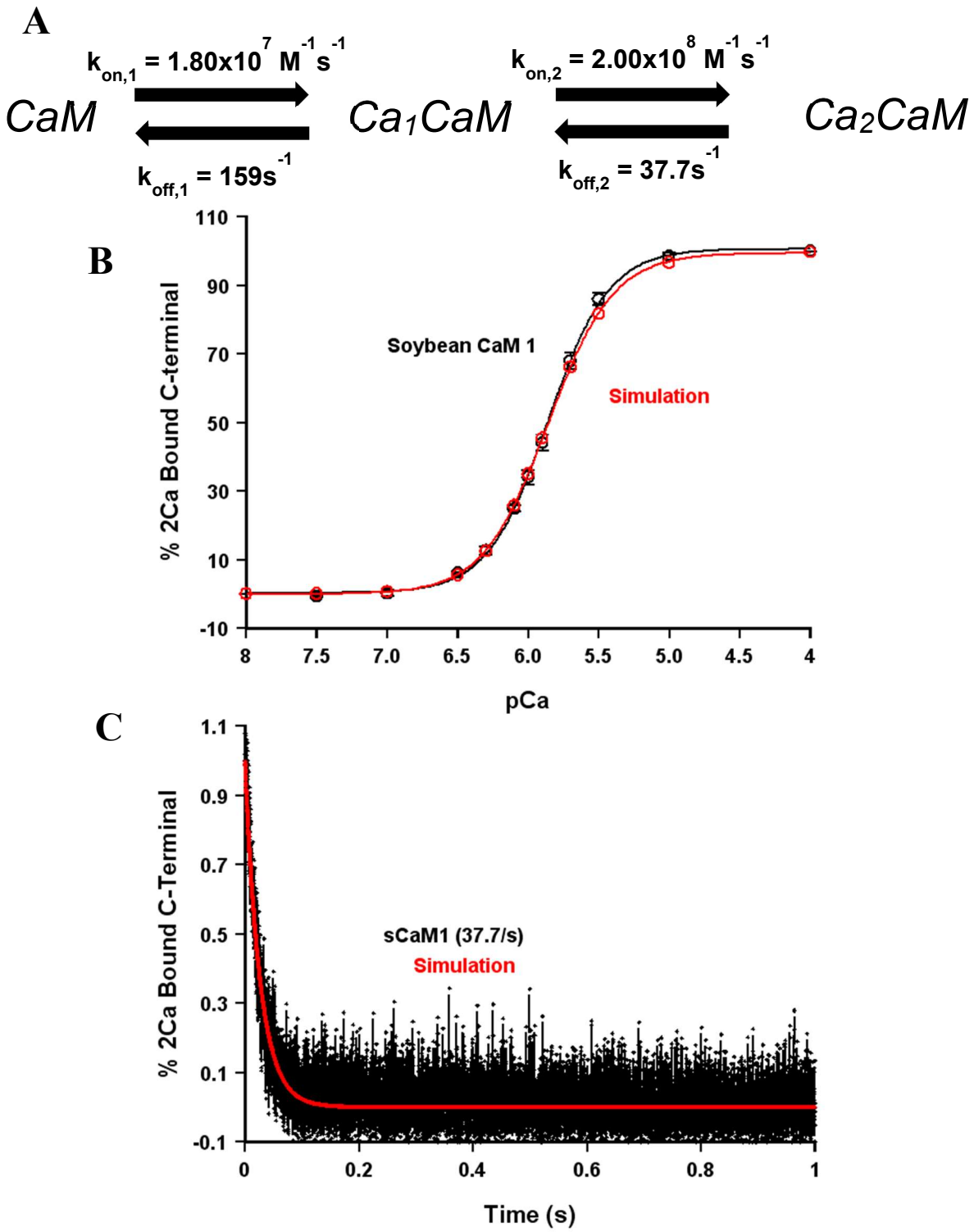


Figure 5. Summary of the model for the C-terminal domain of sCaM1. (a) Schematic of the model with rate parameters, (b) Overlay of real steady-state affinity data with a simulated curve, and (c) Overlay of real calcium dissociation data with a simulated curve.

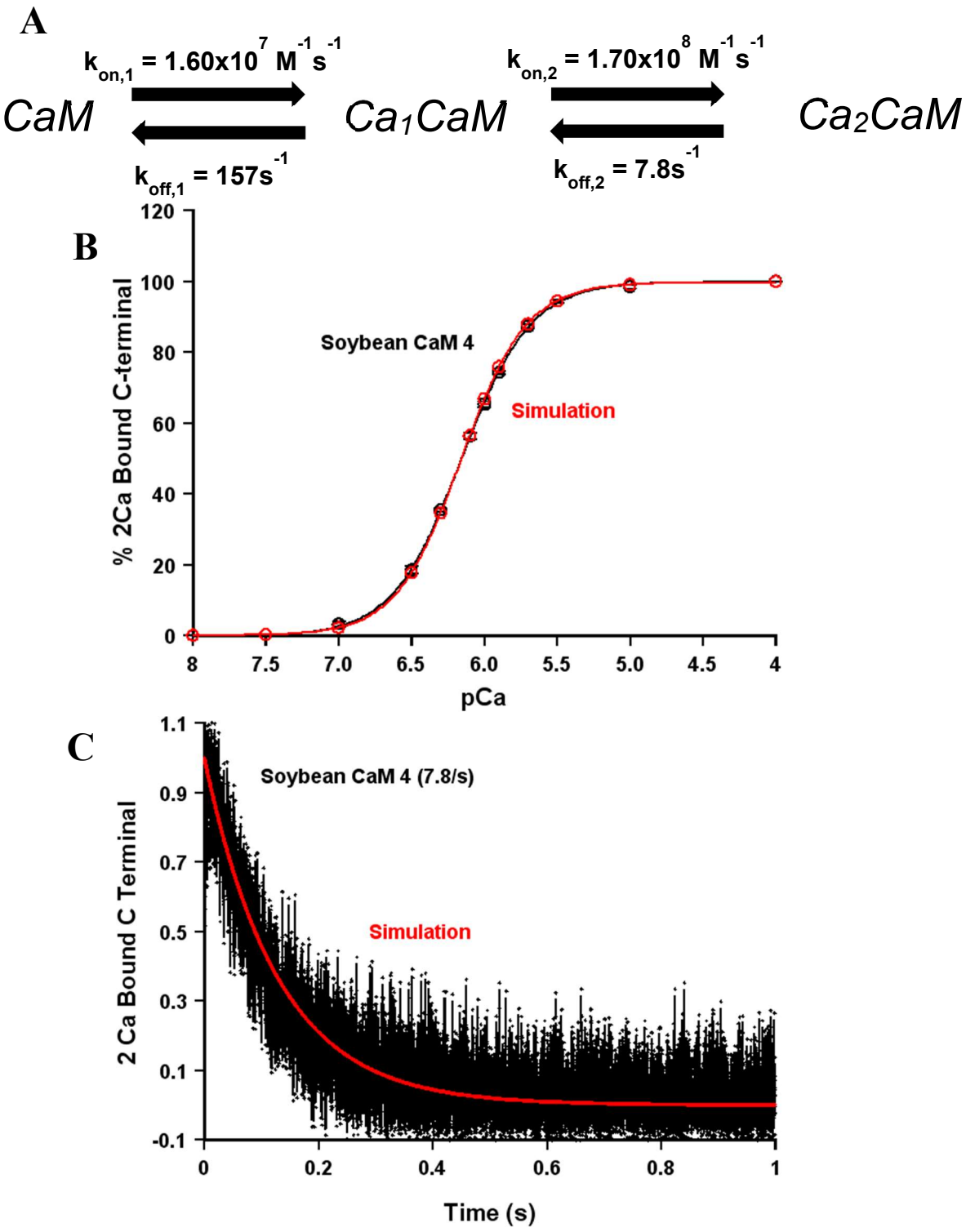


Figure 6. Summary of the model for the C-terminal domain of sCaM4. (a) Schematic of the model with rate parameters, (b) Overlay of real steady-state affinity data with a simulated curve, and (c) Overlay of real calcium dissociation data with a simulated curve.

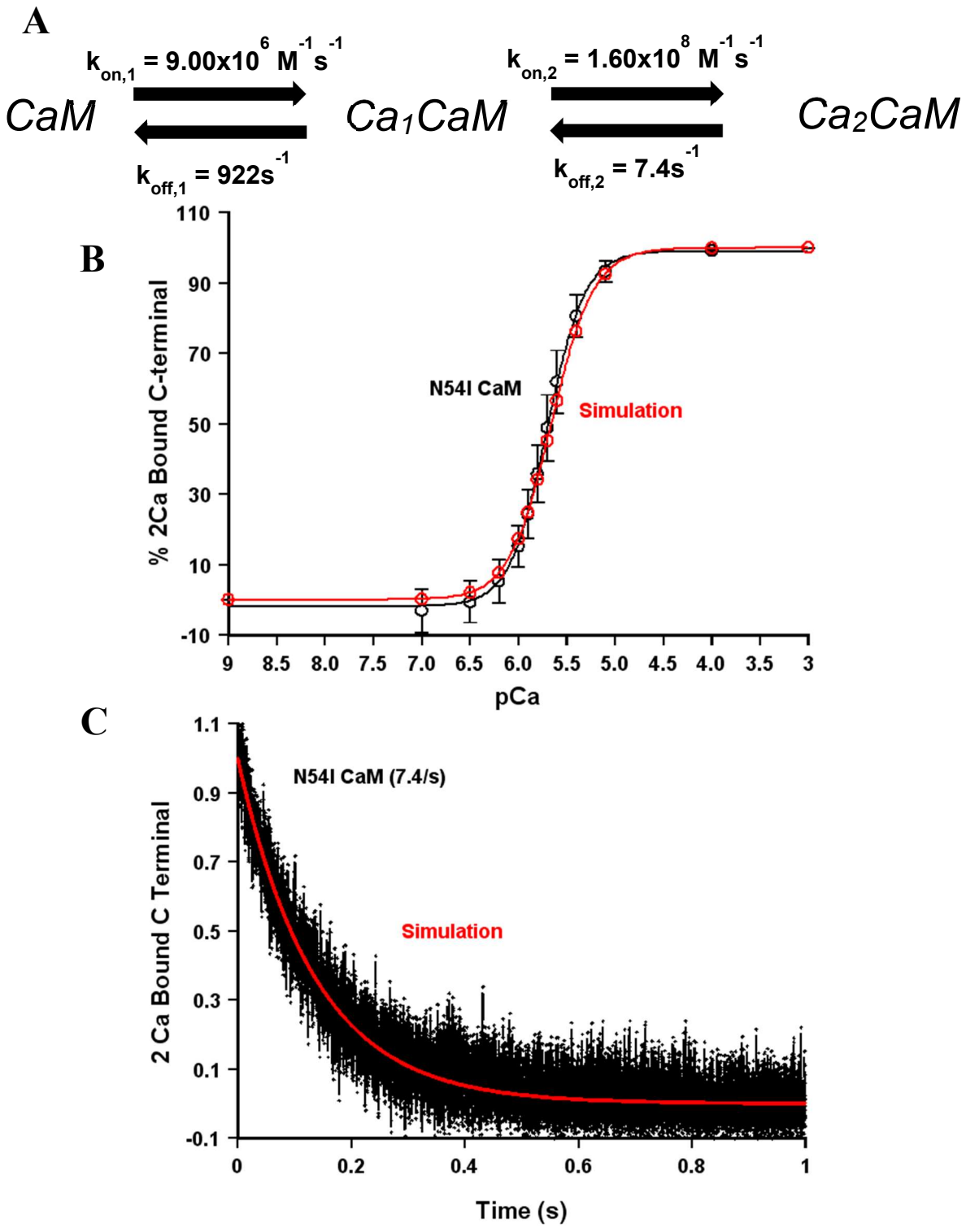


Figure 7. Summary of the model for the C-terminal domain of N54I CaM. (a) Schematic of the model with rate parameters, (b) Overlay of real steady-state affinity data with a simulated curve, and (c) Overlay of real calcium dissociation data with a simulated curve.

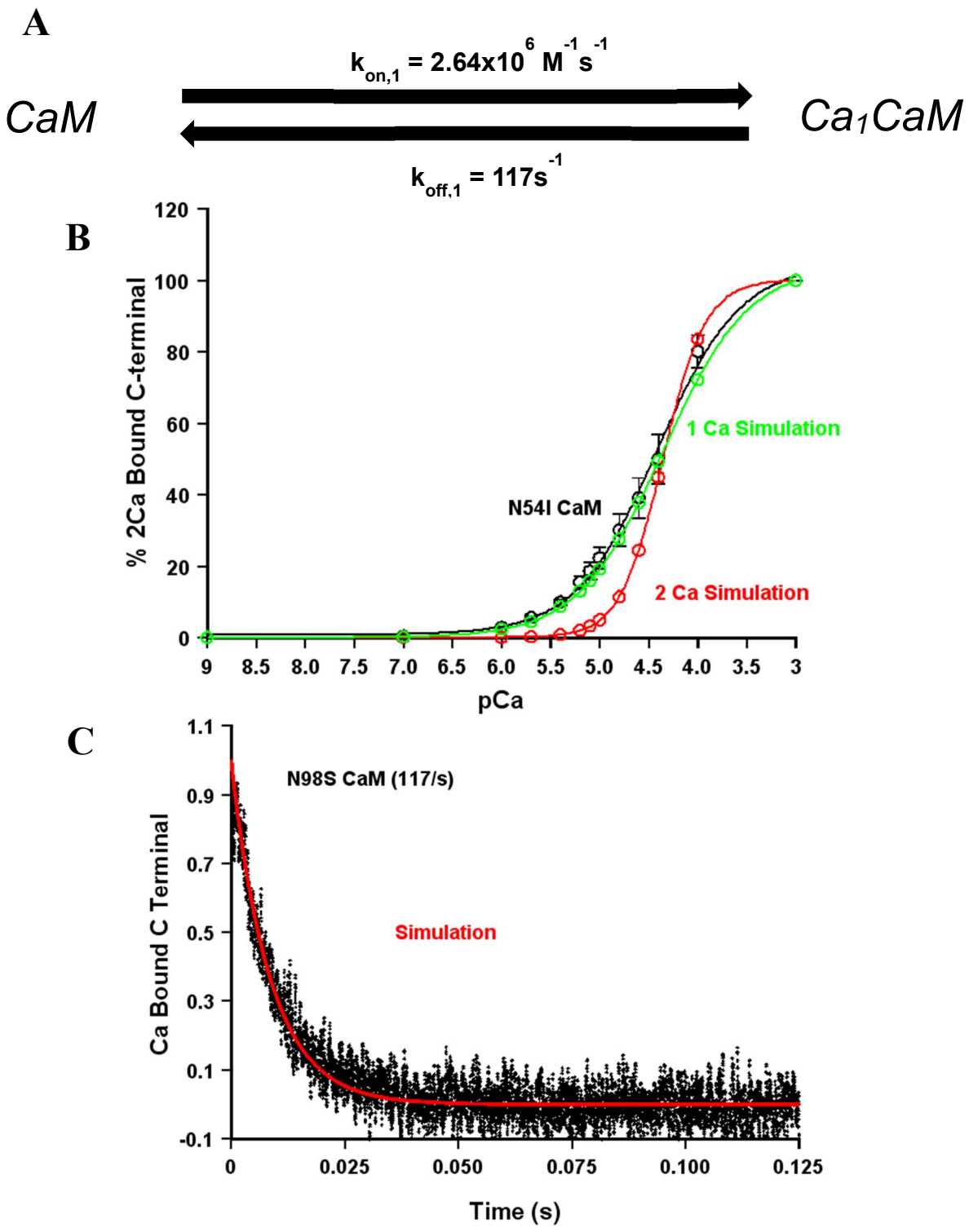


Figure 8. Summary of the model for the C-terminal domain of N98S CaM. (a) Schematic of the model with rate parameters, (b) Overlay of steady-state affinity data with simulated curves based on 2-Ca or 1-Ca models, and (c) Overlay of calcium dissociation data with a simulated curve.

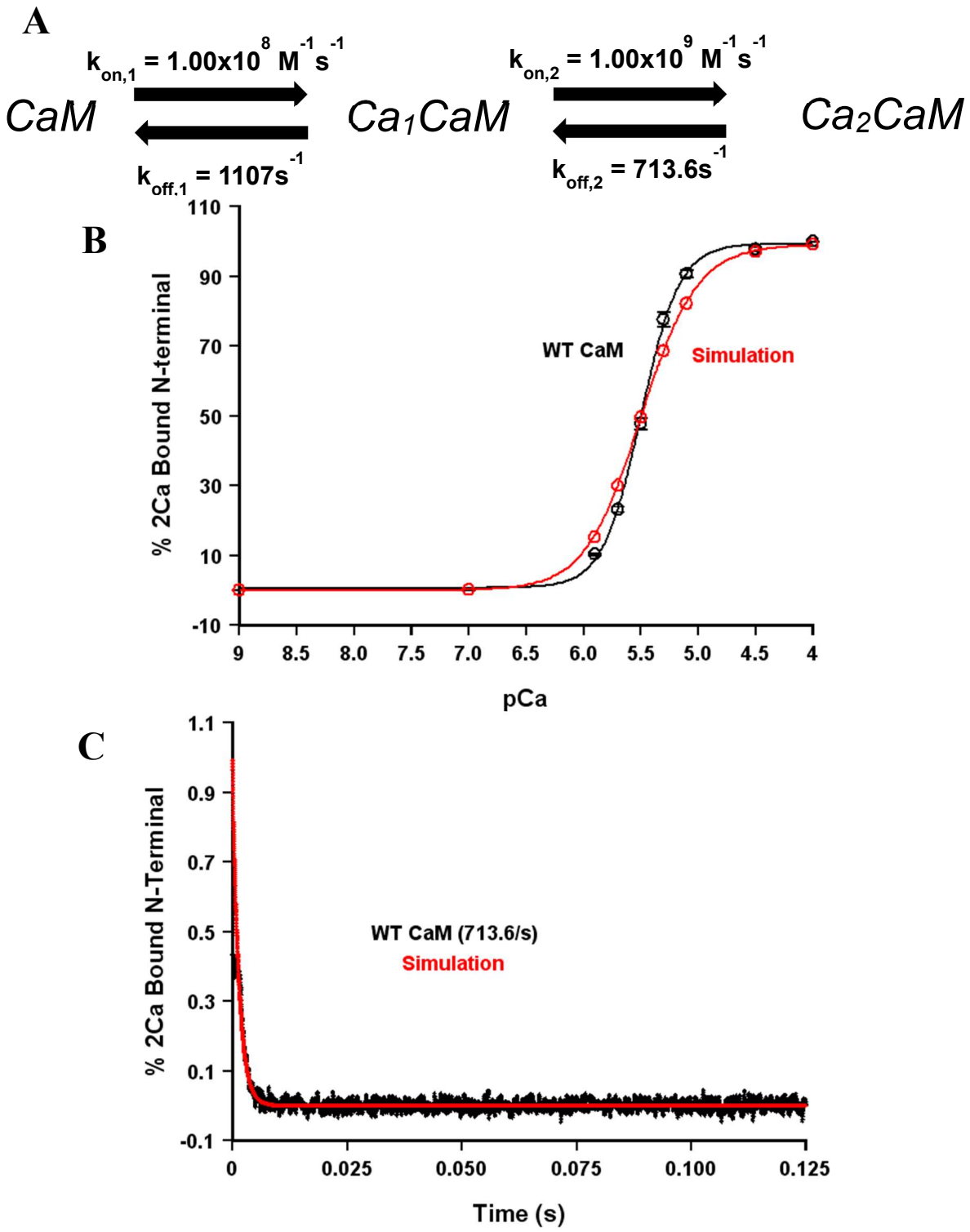


Figure 9. Summary of the model for the N-terminal domain of wild type CaM. (a) Schematic of the model with rate parameters, (b) Overlay of real steady-state affinity data with a simulated curve, and (c) Overlay of real calcium dissociation data with a simulated curve.

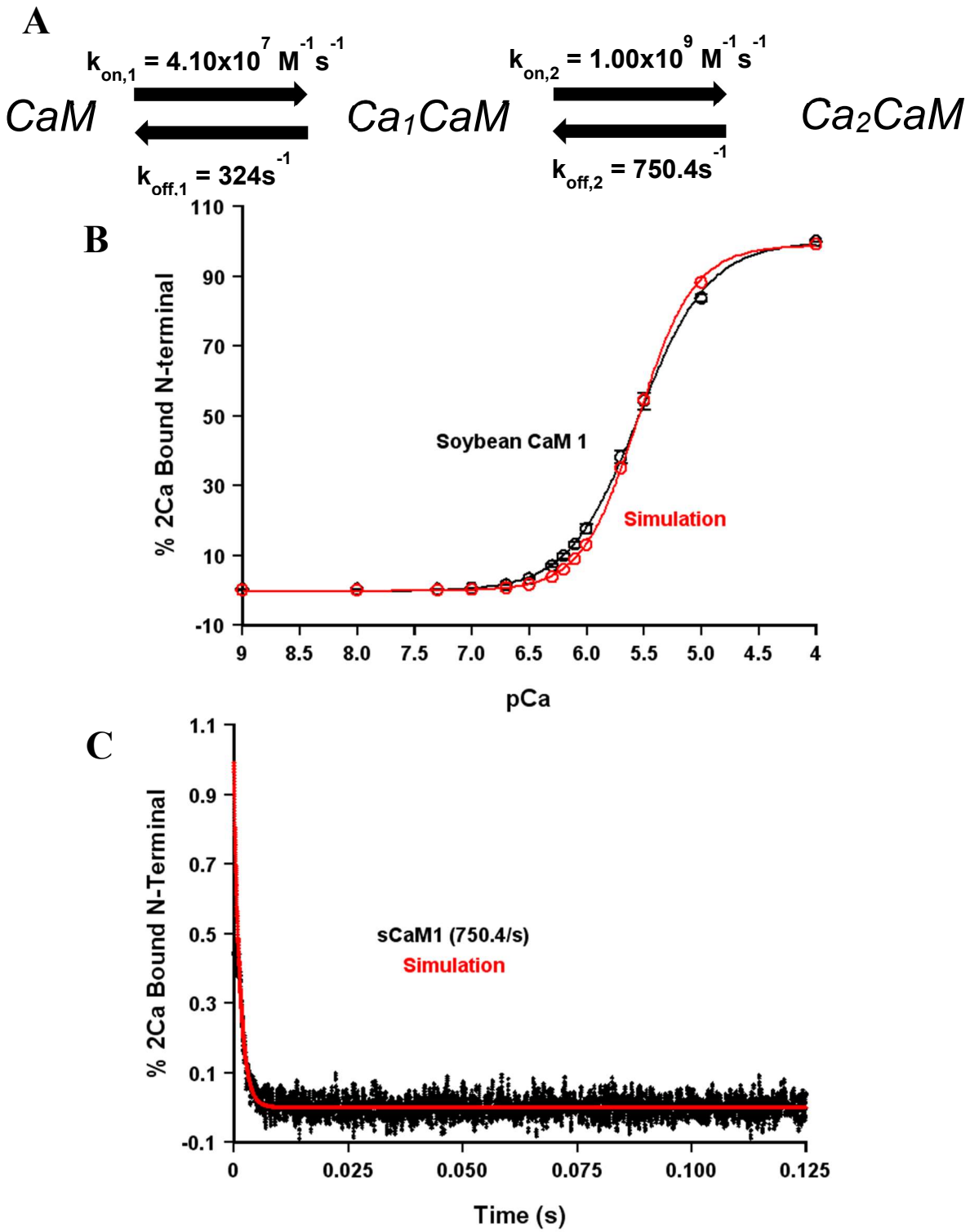


Figure 10. Summary of the model for the N-terminal domain of sCaM1. (a) Schematic of the model with rate parameters, (b) Overlay of real steady-state affinity data with a simulated curve, and (c) Overlay of real calcium dissociation data with a simulated curve.

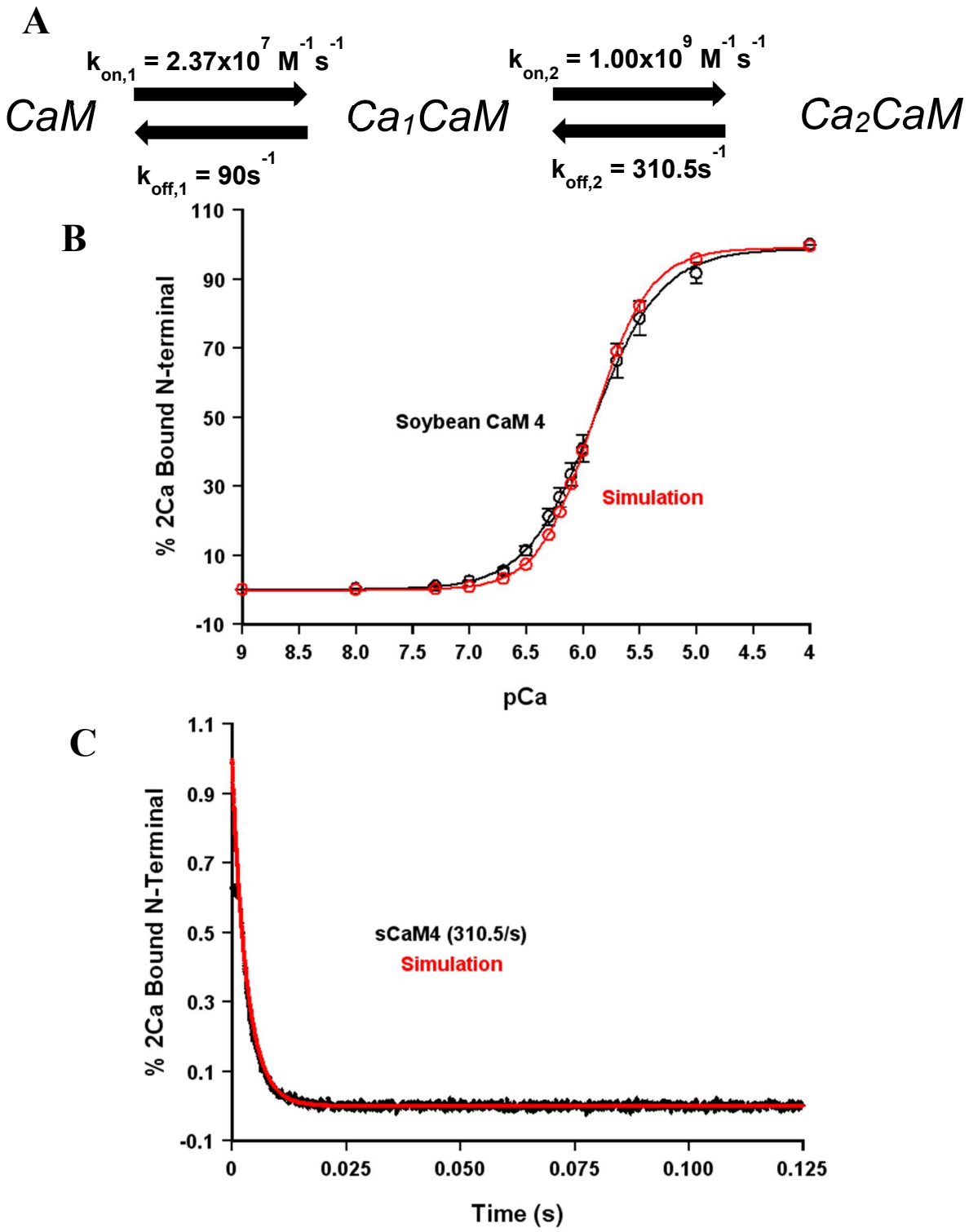


Figure 11. Summary of the model for the N-terminal domain of sCaM4. (a) Schematic of the model with rate parameters, (b) Overlay of real steady-state affinity data with a simulated curve, and (c) Overlay of real calcium dissociation data with a simulated curve.

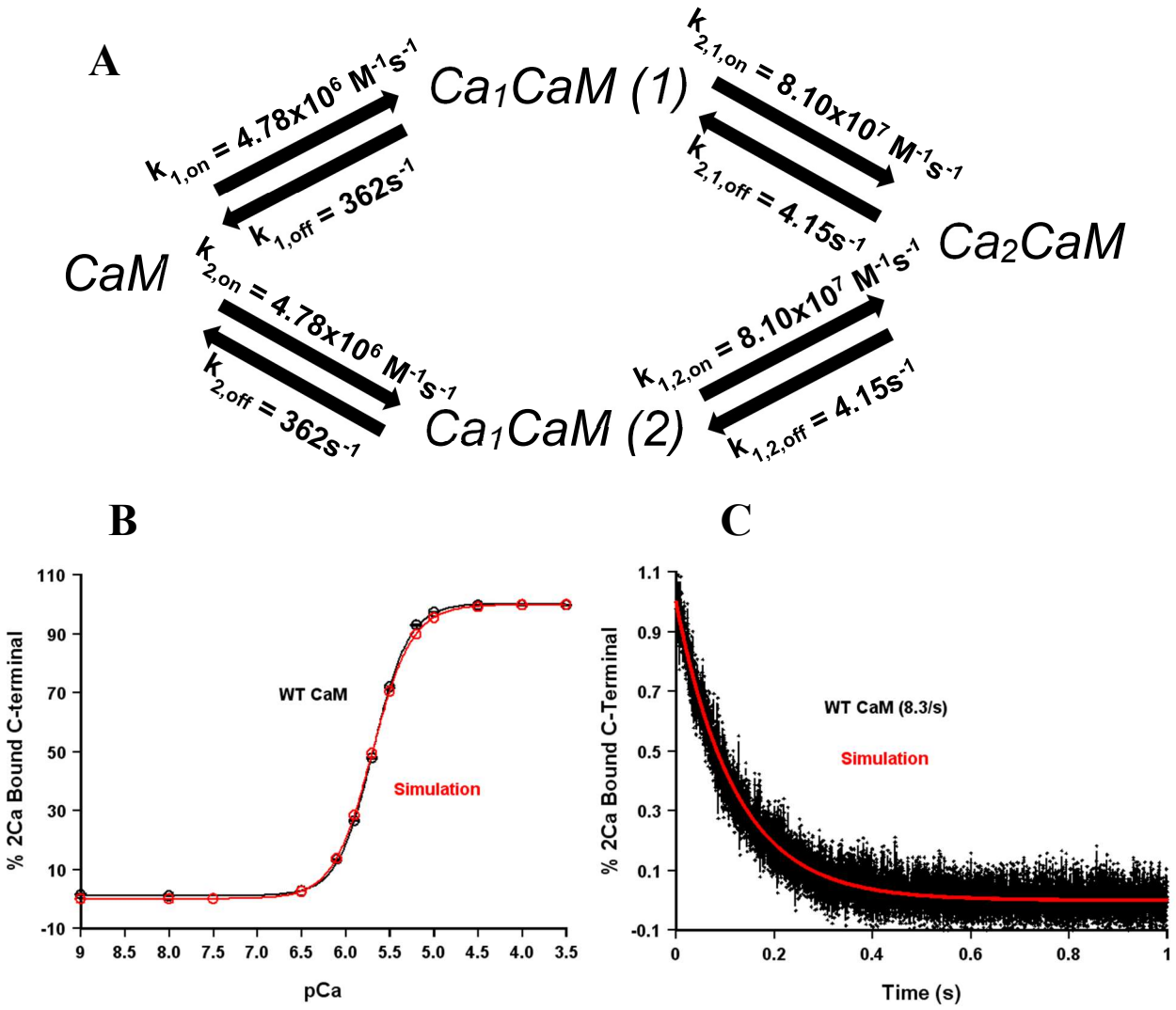
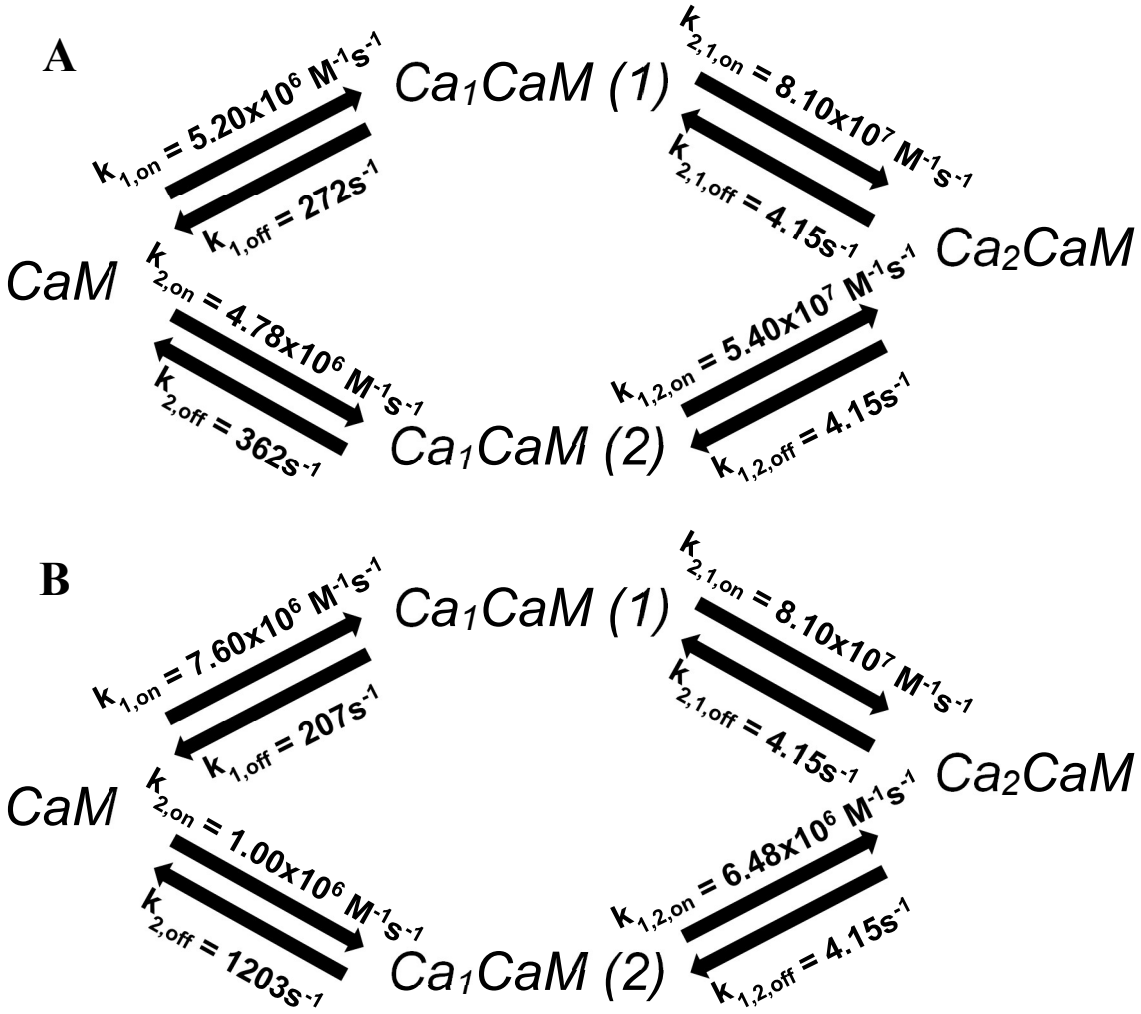
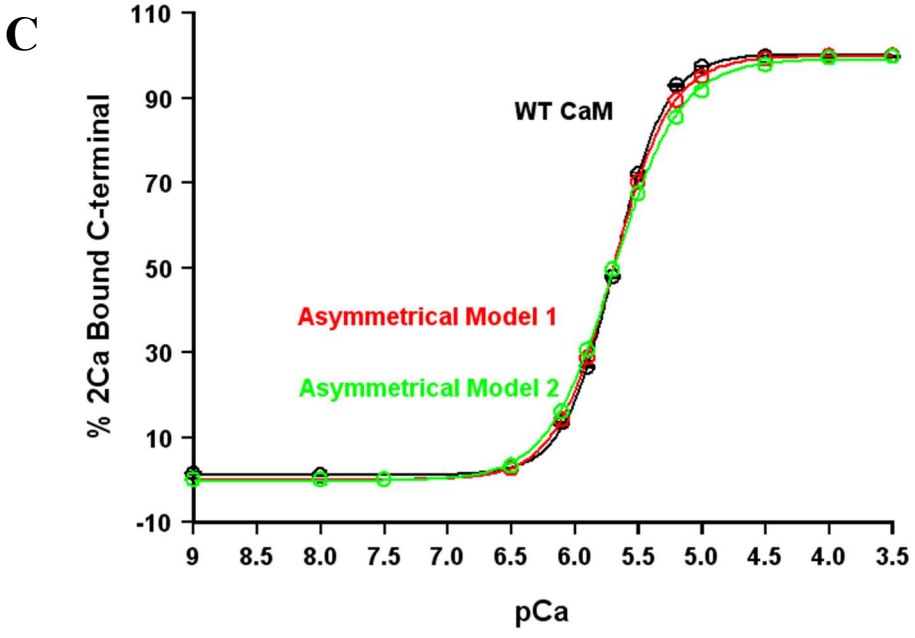


Figure 12. Summary of the symmetrical parallel model for the C-terminal domain of wild type CaM. (a) Schematic of the model with rate parameters, (b) Overlay of steady-state affinity data with a simulated curve, and (c) Overlay of calcium dissociation data with a simulated curve.


A
CaM
Ca1CaM (1)
Ca2CaM
Ca1CaM (2)
k1,on = 5.20x10^6 M^-1s^-1
k1,off = 272s^-1
k2,1,on = 8.10x10^7 M^-1s^-1
k2,1,off = 4.15s^-1
k2,on = 4.78x10^6 M^-1s^-1
k2,off = 362s^-1
k1,2,on = 5.40x10^7 M^-1s^-1
k1,2,off = 4.15s^-1
B
CaM
Ca1CaM (1)
Ca2CaM
Ca1CaM (2)
k1,on = 7.60x10^6 M^-1s^-1
k1,off = 207s^-1
k2,1,on = 8.10x10^7 M^-1s^-1
k2,1,off = 4.15s^-1
k2,on = 1.00x10^6 M^-1s^-1
k2,off = 1203s^-1
k1,2,on = 6.48x10^6 M^-1s^-1
k1,2,off = 4.15s^-1


C
% 2Ca Bound C-terminal
110
90
70
50
30
10
-10
9
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
pCa
WT CaM
Asymmetrical Model 1
Asymmetrical Model 2

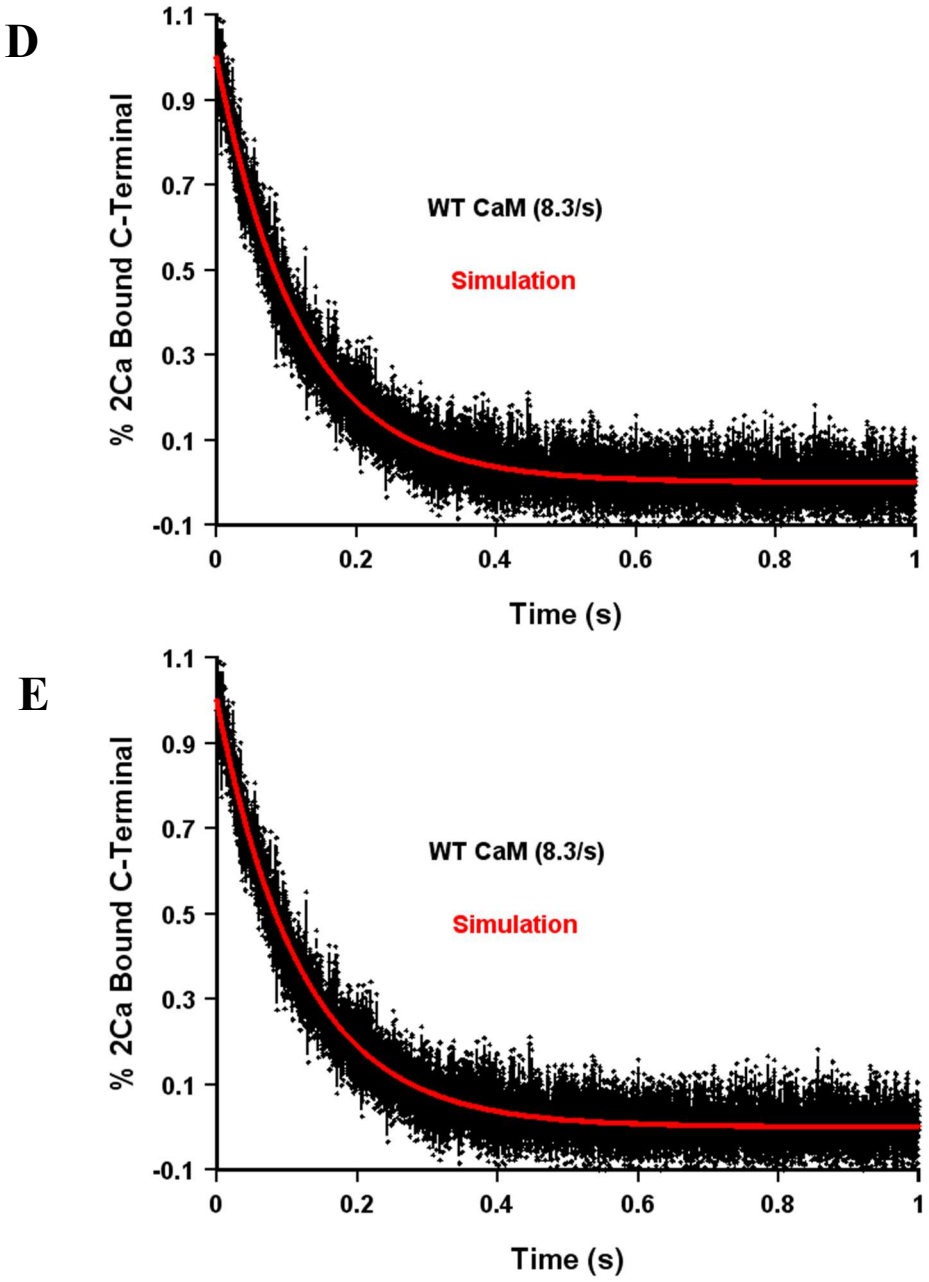


Figure 13. Summary of two asymmetrical parallel models for the C-terminal domain of wild type CaM. (a) Schematic of model one with rate parameters, (b) Schematic of model two with rate parameters, (c) Overlay of real steady-state affinity data with a simulate curve from each model, (d) Overlay of real calcium dissociation data with a simulated curve from model one, and (e) Overlay of real calcium dissociation data with a simulated curve from model two.

Table 1. Summary of the two potential models for the calcium binding properties of slow skeletal troponin.

Model	$k_{on,2}$	$k_{off,2}$	$[TnI]_{eff}$
Cardiac	9.00 x 10^7 $M^{-1}s^{-1}$	110 s^{-1}	12.3μM
1	2.955 x 10^7 $M^{-1}s^{-1}$	110 s^{-1}	26.6μM
2	2.313 x 10^7 $M^{-1}s^{-1}$	85 s^{-1}	8.49μM

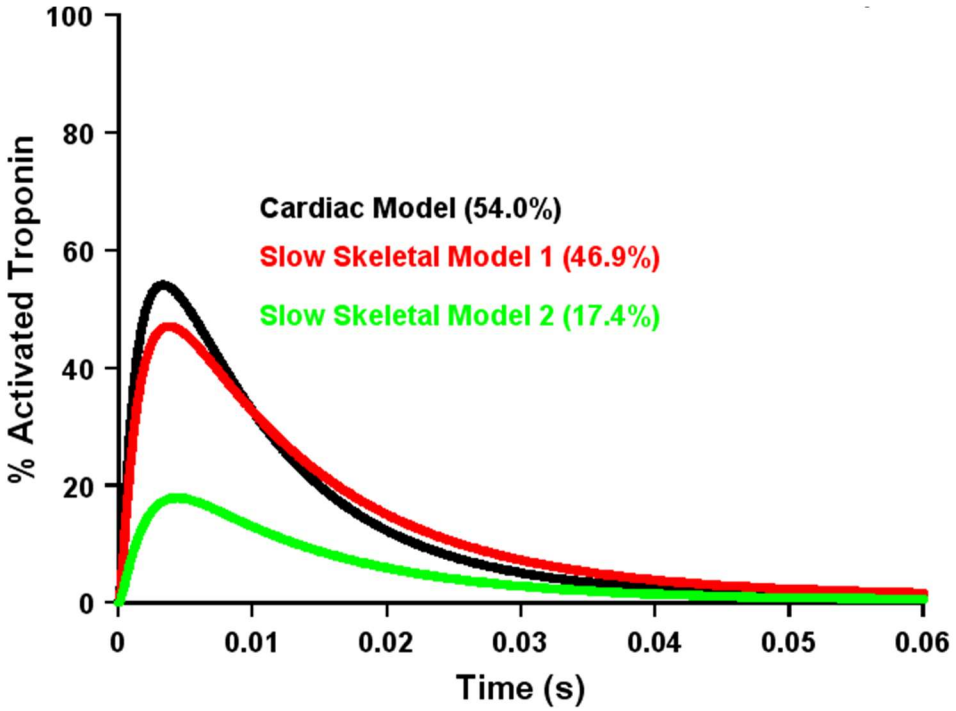


Figure 14. Simulated transient occupancy experiment of the thin filament with different troponin isoforms and models.

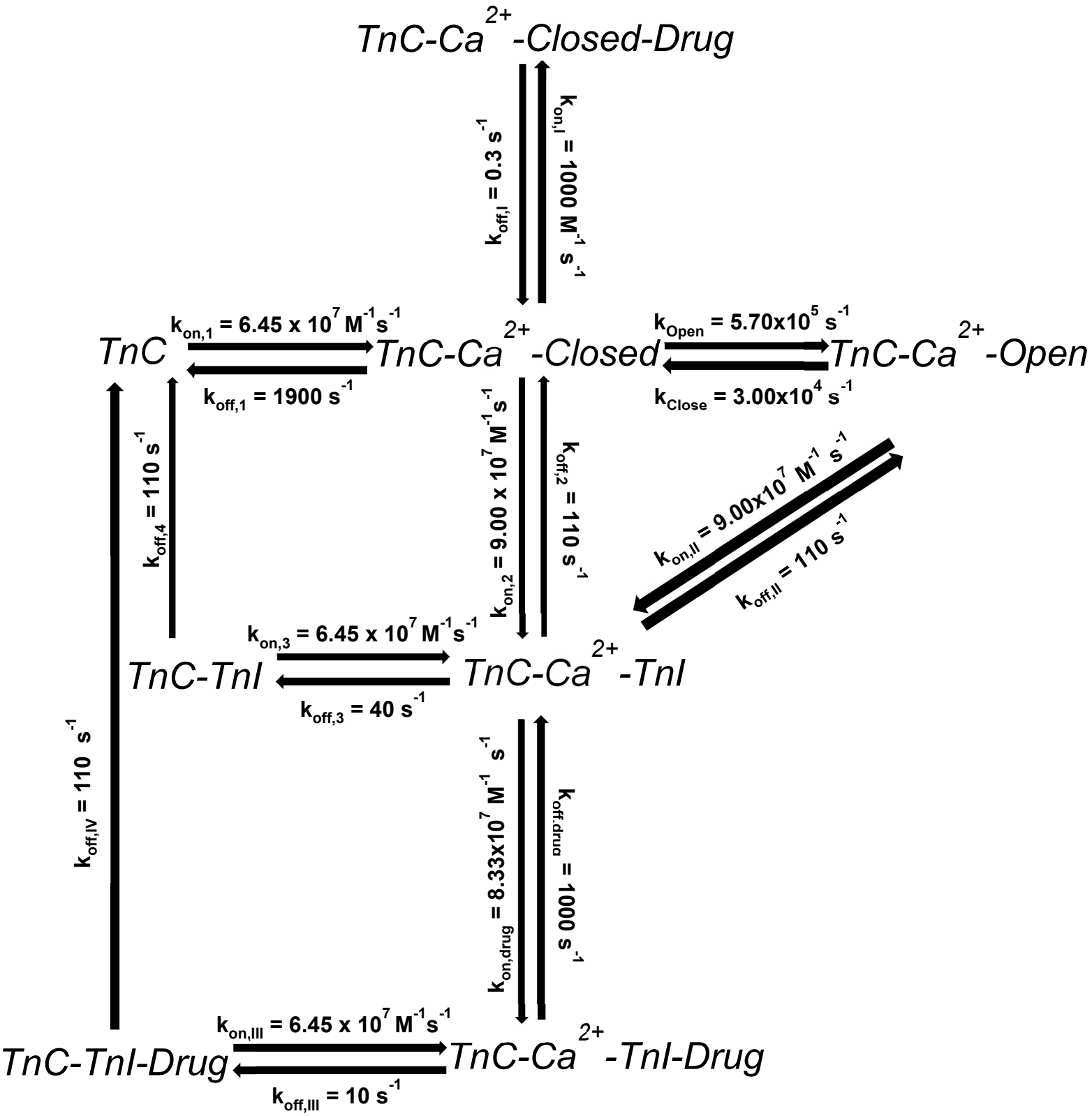


Figure 15. Schematic of the model for the small molecule binding to troponin.

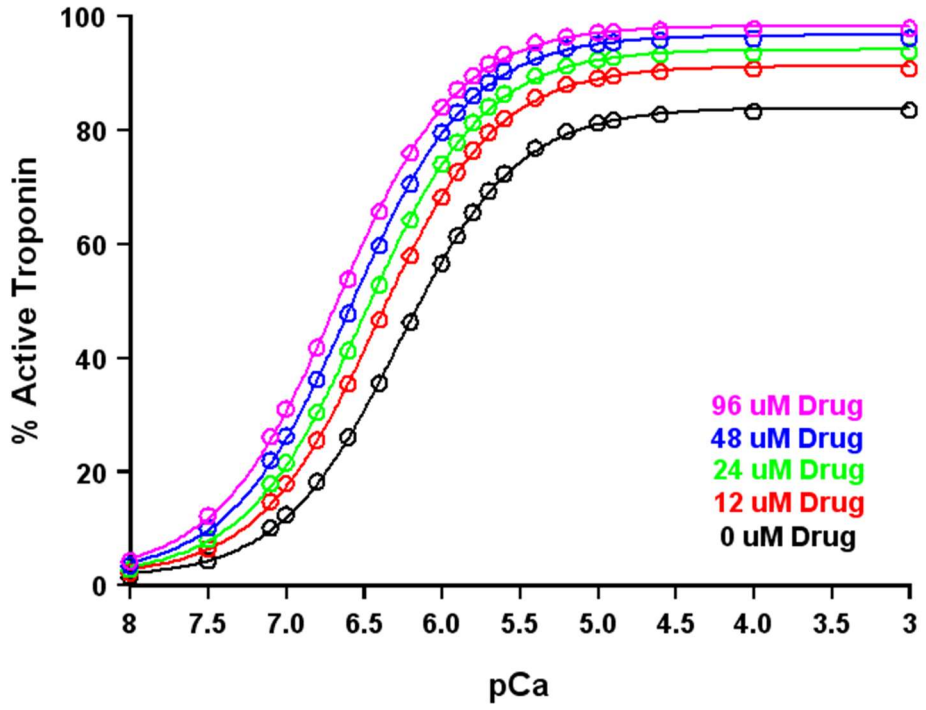


Figure 16. Simulated titration curves of the thin filament with increasing concentrations of the small molecule drug following active troponin.

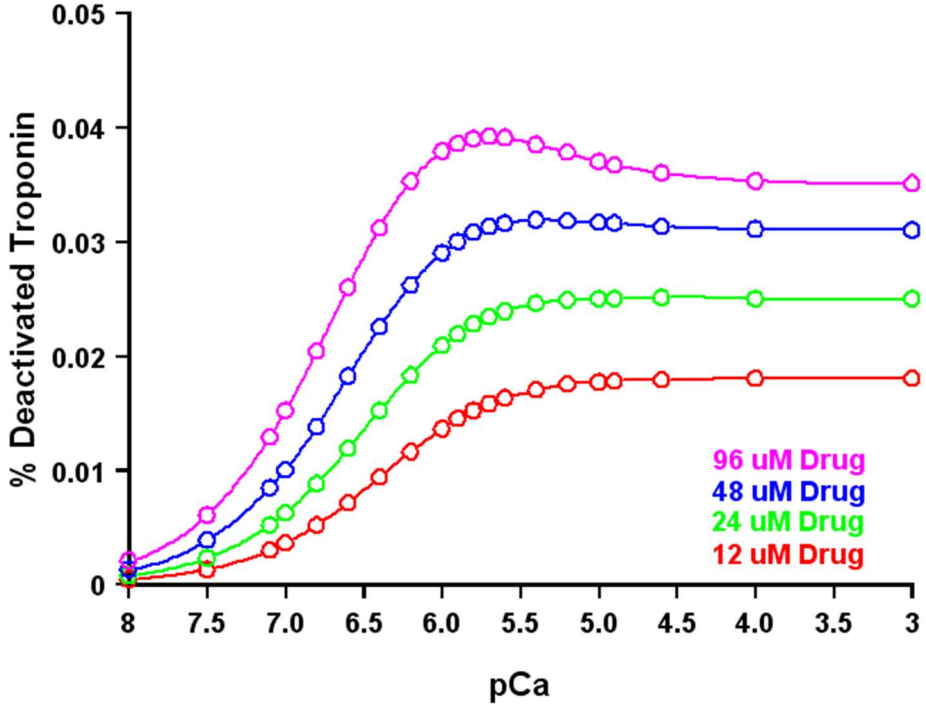


Figure 17. Simulated titration curves of the thin filament with increasing concentrations of the small molecule drug following the deactivated troponin species.

Table 2. Concentration of the small molecule drug required for the maximum activation of troponin at saturating calcium to be 50%, based on the K_D of the inhibitory step (binding of the drug to the TnC-Ca^{2+}-Closed state), in the thin filament (Low [TnI]) and the troponin complex (High [TnI]).

Inhibitory K_D	Low [TnI]	High [TnI]
300 μM	376.2 M	9829 M
30 μM	3.82 M	99.9 M
3 μM	368 mM	9.98 M
300 nM	22.9 mM	990 mM
30 nM	3.68 μM	90.1 mM
3 nM	658 nM	581 μM
300 pM	424 nM	1.95 μM
30 pM	403 nM	631 nM